# Digital Color Image Processing and Psychophysics Within the Framework of a Human Visual Model

OLIVIER DOMINIQUE FAUGERAS

UNIVERSITY OF UTAH

(5)









JUNE 1976
UTEC-CSc-77-029
COMPUTER SCIENCE, UNIVERSITY OF UTAH
SALT LAKE CITY, UTAH 84112

(6) DIGITAL COLOR IMAGE PROCESSING AND PSYCHOPHYSICS WITHIN THE FRAMEWORK OF A HUMAN VISUAL MODEL.

by

(10) Olivier Dominique Faugeras

(9) Technical rept.,





(12) 133p.



(11) June 1976

(14) UTEC-CSc-77-029

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract (15) No. DAHC15-73-C-0363, ARPA Order-2477

TABLE OF CONTENTS

## LIST OF FIGURES

LIST OF TABLES

## ABSTRACT

A three-dimensional homomorphic model of human color vision based on neurophysiological and psychophysical evidence is presented. This model permits the quantitative definition of perceptually important parameters such as brightness, saturation, hue and strength. By modelling neural interaction in the human visual system as three linear filters operating on perceptual quantities, this model accounts for the automatic gain control properties of the eye and for brightness and color contrast effects.

In relation to color contrast effects, a psychophysical experiment was performed using a high quality color television monitor driven by a general purpose digital computer. This experiment, based on the cancellation by human subjects of simultaneous color contrast illusions, allowed the measurement of the low spatial frequency part of the frequency responses of the filters operating on the two chromatic channels of the human visual system. The experiment is described and its results are discussed.

Next, the model is shown to provide a suitable framework in which to perform digital images processing tasks. First, applications to color image enhancement are presented and discussed in relation to photographic masking techniques and to the handling of digital color images. Second, application of the model to the definition of a distortion measure between color images (in the

sense of Shannon's rate-distortion theory), meaningful in terms of human evaluation, is shown. Mathematical norms in the "perceptual" space defined by the model are used to evaluate quantitatively the amount of subjective distortion present in artificially distorted color images. Third, applications to image transmission and coding are presented. Results of a coding experiment yielding digital color images coded at an average bit rate of 1 bit/pixel are shown.

Finally conclusions are drawn about the implications of this research from the standpoint of psychophysics and of digital image processing.

# CHAPTER 1

## PSYCHOPHYSICAL AND NEUROPHYSIOLOGICAL BACKGROUND

### 1.1 Colorimetry

The main fact about color vision is its basic trichromatic nature. Overwhelming support has been provided over the past one hundred and fifty years for the notion first put forth clearly by Thomas Young that three separate varieties of receptors underlie color vision.

In particular the experimental laws of color-matching are an indirect verification of this fact. They state that over a wide range of conditions of observation colors can be matched by additive mixtures of three fixed primary colors. These primaries must be such that none of them can be matched by a mixture of the other two. Furthermore those matches are linear over a wide range of observing conditions which implies two things. First, the match between any two colors continues to hold if the corresponding stimuli are increased or reduced by the same amount, their respective relative spectral energy distributions being unchanged (scalability condition). Second, if colors C and D match and colors E and F match, then the additive mixture of colors C and E matches the corresponding additive mixture of colors D and F.

This linearity property allows us to represent colors by three-dimensional vectors and color matches by linear equations

Stockham [50, 51], was the first one to suggest that a multiplicative rather than linear system was better fitted to the structure of black and white images and furthermore was similar to the processing occuring in the earlier stages of the human visual path. His efforts to define and use a multiplicative homomorphic visual model have been carried on and fully justified by Baudelaire [1] and extended by Baxter [2]. The use of the model for the definition of a subjective distortion measure has been explored by Mannos [33] and Rom [39].

Very briefly, the multiplicative homomorphic model has the following property: by modelling the nonlinearity in the absorption of light energy in the retina as well as neural interaction later on the visual path, it accounts for brightness contrast and brightness constancy and realizes the separation of reflectance from illumination.

Effects similar to those previously mentioned are also present in color vision. Color constancy effects which show the tremendous ability of the eye to adapt instantaneously to ambiant illumination even if heavily color biased are well known [6, 32]. Color contrast effects and color Mach Bands have also been recognized [22, 28].

But a clear formalization of these phenomena is hampered by the three-dimensionality of color vision. A color can be defined in terms of its brightness, hue and saturation and interaction between colors can cause the change of all these parameters. Clearly, a model of color adaptation and a quantitative description of such parameters as color brightness and saturation would be welcome.

Knowing this, it seems very appealing to attempt to model human

color vision using a multiplicative homomorphic model. This approach should guarantee that the new model reduces to the old one for achromatic vision. Also, it would allow us to use the powerful tool of Fourier analysis which has proven to be well suited to a precise description of achromatic brightness contrast and constancy effects [1, 6] and is therefore hoped to perform as well for color effects.

In our opinion, the achievements of this research are twofold: On the side of psychophysics, we have shown that by using a model based on neurophysiological and psychophysical evidence as a theoretical background, and a high precision television monitor driven by a digital computer as a tool, we could obtain some relevant experimental results about the way chromatic information is processed by the human visual system.

On the side of image processing, we have shown that this model allowed us to think about and perform in a meaningful way our image processing tasks.

Chapter 1 lays the psychophysical and neurophysiological basis which is used in chapter 2 to develop a quantitative description of achromatic and chromatic information processing in the visual path. Chapter 3 describes the psychophysical experiment by which we probed the chromatic channels. Chapters 4 and 5 are dedicated to the applications of the model to image processing. Chapter 4 describes some ways of enhancing color images. Chapter 5 elaborates on the use of the model to define a perceptual distance between color images and on related problems of transmission and coding. Chapter 6 draws the conclusions of this research project.

# INTRODUCTION

The research presented in this report is concerned with image processing. In a very general sense, image processing can be thought of as a transformation of an image from one form into another. The results can appear as another image, a set of decisions, a model or parametrization. Here, we will restrict ourselves to the case where the transformation yields as its output another image which is different from the original in some desirable way.

Since a human observer is likely to be the last element in the processing chain, it seems natural to become interested in the properties of the human visual system. The idea of applying the current knowledge about the human visual system to the field of image processing is a recent one though. It stemmed from the availability of digital methods on one hand and the lack of a quantitative definition of subjective image quality on the other hand.

Indeed, digital methods are characterized by their flexibility and their precision. Good subjective distortion measures built from a clear and concise formalism have begun to emerge for achromatic images. For color the knowledge is not as advanced. In both cases, still a lot remains to be done in order to further sharpen our understanding of the way visual information is processed and extracted by the visual neural system.

between such vectors. If C represents a given color and R, G and B represent unit amounts of three fixed primaries then the equation

$$C=RR+GG+BB$$

expresses the fact that the given color is matched by an additive mixture of quantities R, G and B respectively of the given primaries. R, G and B are called the tristimulus values of the given color in the particular set of primaries used. Another set of frequently used quantities is the chromaticity coordinates r, g, b defined by

$$r = R/(R+G+B)$$

$$g = G/(R+G+B)$$

$$b = B/(R+G+B)$$

The linearity property also allows us to define completely the color-matching properties of the observing eye in the given primary system by three functions of wavelength $r(\lambda)$, $g(\lambda)$ and $b(\lambda)$ called color-matching functions. If the color C has a spectral energy distribution $p(\lambda)$ then its tristimulus values are

$$R=\int p(\lambda)r(\lambda)d\lambda$$

$$G=\int p(\lambda)g(\lambda)d\lambda$$

$$B=\int p(\lambda)b(\lambda)d\lambda$$

where the integrals are taken over the visible spectrum. Colors $C_1$ and $C_2$ of spectral energy distributions $p_1(\lambda)$ and $p_2(\lambda)$ match if and only if

$$\int p_1(\lambda)r(\lambda)d\lambda=\int p_2(\lambda)r(\lambda)d\lambda \qquad (1.1a)$$

$$\int p_1(\lambda)g(\lambda)d\lambda=\int p_2(\lambda)g(\lambda)d\lambda \qquad (1.1b)$$

$$\int p_1(\lambda)b(\lambda)d\lambda=\int p_2(\lambda)b(\lambda)d\lambda \qquad (1.1c)$$

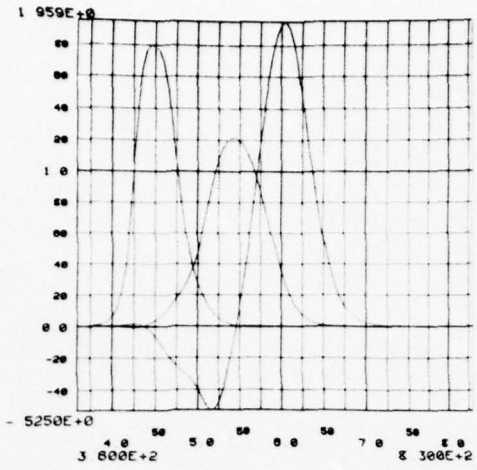



Fig. 1.1- A set of color-matching functions for the primary system of monochromatic lights at 700 nm, 546 nm, 436 nm.

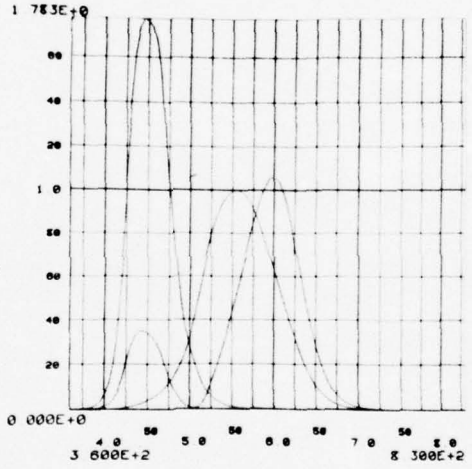



Fig. 1.2- The 1931 C.I.E. color-matching functions $x(\lambda)$, $y(\lambda)$, $z(\lambda)$.

A set of such color-matching functions is shown in figure 1.1.

The linearity property also allows us to describe a change of primaries as a simple matrix multiplication operation whereby the new tristimulus values L, M and S of a given color and the new color-matching functions $l(\lambda)$, $m(\lambda)$ and $s(\lambda)$ are given by

$$\vec{T}_n = A\,\vec{T}_o \qquad \vec{\tau}_n = A\vec{\tau}_o$$

where $\vec{T}_n = [L,M,S]^t$, $\vec{T}_o = [R,G,B]^t$, $\vec{\tau}_n = [l(\lambda),m(\lambda),s(\lambda)]^t$, $\vec{\tau}_o = [r(\lambda),g(\lambda),b(\lambda)]^t$ and A is a three by three matrix whose column vectors are the tristimulus values of the old primaries in the system of new primaries.

A standard observer has been developed in 1931 by the CIE ("Commission Internationale de l'Eclairage") whose color-matching functions are denoted by $x(\lambda)$, $y(\lambda)$ and $z(\lambda)$ and shown in figure 1.2. The matches predicted with their aid using equations 1.1a to 1.1c represent average matches for humans with normal color vision as opposed to humans with any one of the several forms of color blind vision.

## 1.2 Spectral sensitivities of the cone receptors

The physiological basis of trichromatic color matches is the linear absorption of light at photopic levels of illumination (when the effect of rods receptors is negligible) by three types of cone receptors in the human fovea. However, specification of the spectral absorption characteristics of the three different cone types has been difficult since it has not been possible to extract cone pigments from the eyes of primates [12]. A number of different approaches has been taken to yield this information which is

essential to an understanding of color vision. Psychophysical evidence has come from Willmer [60], Blackwell and Blackwell [3], Stiles [49] and Pitt [36] among many. Rushton [40, 41] has attempted to answer the same question by reflection densitometry. Attempts have also been made to measure those curves from unit recordings later in the visual pathway (De Valois 1965, [11], Wiesel and Hubel 1966, [59]).

All these studies agree well on two points. The first is that there are three types of cones with their absorption curves peaking at about 445, 540 and 570 nm. The second is that these absorption curves are quite broad and fit to a first approximation the Dartnall nomogram (Dartnall 1953, [7]).

The common reference to those receptors as "red", "green" and "blue" receptors stems from the spread of Young-Helmholtz type of theories of color-vision. This terminology is confusing. For example, let us consider the case of the 570 pigment cone so misleadingly called by many the "red receptor". Not only is it not limited in its sensitivity to the part of the spectrum we see as red (about 600 to 700 nm) but it is not even maximally sensitive in this spectral region. In fact it is more sensitive to that part of the spectrum we see as pure green (510 nm) than to pure red (650 nm)!

So both on account of its point of maximal absorption and of the breadth of its absorption curve it is a gross misnomer to call the 570 receptor a "red receptor" or the 540 cone a "green receptor". But worst is the fact that calling those receptors "red" or "green" implies that their excitation signals "red" or "green" further away on the visual path.

We will use the terms L-cones (containing the long-wavelength peaking pigment), M-cones (containing the medium-wavelength peaking pigment) and S-cones (containing the short-wavelength peaking pigment) to denote the cone types containing the 570, 540 and 445 nm peaking pigments.

## 1.3 Nonlinear response of the cones

Physiological studies have shown that the activity or response of the cone receptors is proportional not to the intensity of the stimulus but rather to its logarithm over a large range of stimulus intensities [54, 37, 6]. One psychophysical equivalent of these physiological data is the well known Fechner's law which states that the just noticeable brightness difference is proportional to the logarithm of the stimulus intensity [27].

It is interesting to point out that this quasi-logarithmic mapping of light intensity into neural activity at the output of the three cones allows the retina to accept inputs over a very wide range of intensities [24]. It has already been recognized [58] and it will be shown quantitatively in this report that this ability for visual stimuli compression at the receptor level together with neural interaction later on the visual path is the fundamental basis for visual adaptation.

## 1.4 The processing of brightness and color information past the retina

We know that in higher vertebrates the retina projects itself to the lateral geniculate nucleus (LGN) and from there to the striate cortex [12]. Unfortunately it is far from possible at our

present state of knowledge to describe the central neural processing of visual information past the LGN. However from the close parallels one finds between the responses of cells at the LGN level and color perception, it would appear that much of the processing of color information takes place before the cortex. With the risk of oversimplifying we will summarize the results obtained by neurophysiological experiments on LGN cells of primates [10, 13, 59].

Four varieties of so-called spectrally opponent cells were found in the LGN. They showed excitation to some wavelengths and inhibition to others when the eye was stimulated with flashes of monochromatic lights. Moreover this type of response was maintained over wide intensity ranges thus indicating that the cells responded differentially to color rather than to luminance differences. The cells whose responses showed maximum excitation around 500 nm and maximum inhibition around 630 nm were called green excitatory, red inhibitory cells (+G-R). Other cells showing rough mirror image responses were called red excitatory, green inhibitory cells (+R-G). The cells showing maximum excitation (inhibition) at 600 nm and maximum inhibition (excitation) at 440 nm were named yellow excitatory, blue inhibitory cells (+Y-B) and blue excitatory, yellow inhibitory cells (+B-Y) respectively.

In addition to these four varieties of spectrally opponent cells, two other classes of cells were found that did not give spectrally opponent responses but rather responded in the same direction (either excitation or inhibition) to lights of all wavelengths. Those cells thus appear not to be involved with color

vision and were named white excitatory, black inhibitory cells (+Wh-Bl) and black excitatory, white inhibitory cells (+Bl-Wh).

Furthermore these studies were able to answer the question of what are the cone inputs to the LGN cells. Red-green cells (+R-G and +G-R) receive their inputs from the L and M-cones. One of the cone type is excitatory, the other inhibitory. Which cones feed into the yellow-blue system is a matter of dispute. It is certain that they receive inputs from the S-cones but whether the other inputs are from the L or M-cones (or possibly both) is uncertain. Since the weight of the evidence seems to indicate that the first situation is the one which actually exists [12], we will assume that the second type of inputs is received from the L-cones. Like the red-green cells, one cone input to the yellow-blue cells is excitatory, the other inhibitory. For the non-opponent cells there is good evidence that both the L and M-cones (and possibly S-cones as well) are feeding into them and acting in the same direction.

There is thus much evidence from unit recording experiments in favor of the existence of separate chromatic and achromatic channels each gaining its information from the same receptors but processing it in different ways. Spectrally non-opponent cells which correspond to the achromatic channel add together the (Log) outputs of the L, M and S-cones. Spectrally opponent cells which correspond to the chromatic channels subtract the (Log) outputs of the L and M-cones for the red-green system, of the L and S-cones for the yellow-blue system.

It is clear that, although the physiological processes involved are quite different from those Hering [25] proposed in 1878, the

cell types found in the LGN correspond very closely to the types of retinal outputs he hypothesized to account for color appearances.

## 1.5 Spatial organization of LGN cells

The cone inputs to the LGN cells are spatially distributed. The spectrally non-opponent cells have a center-surround organization where the center is excitatory and the surround inhibitory or vice versa [59]. With regard to the spectrally opponent cells, the situation is somewhat more complicated. Two types of organization are found. The most common variety has an excitatory input from one cone type in the center and an inhibitory input from the other type in the surround or vice versa. By far the largest number of these cells are red-green cells with only a few yellow-blue cells. The second variety is found much less frequently and consists in spectrally opponent cells in which the two cone inputs have identical spatial distributions. About half of those cells are red-green cells and half yellow-blue cells. This situation is summarized in figure 1.3.

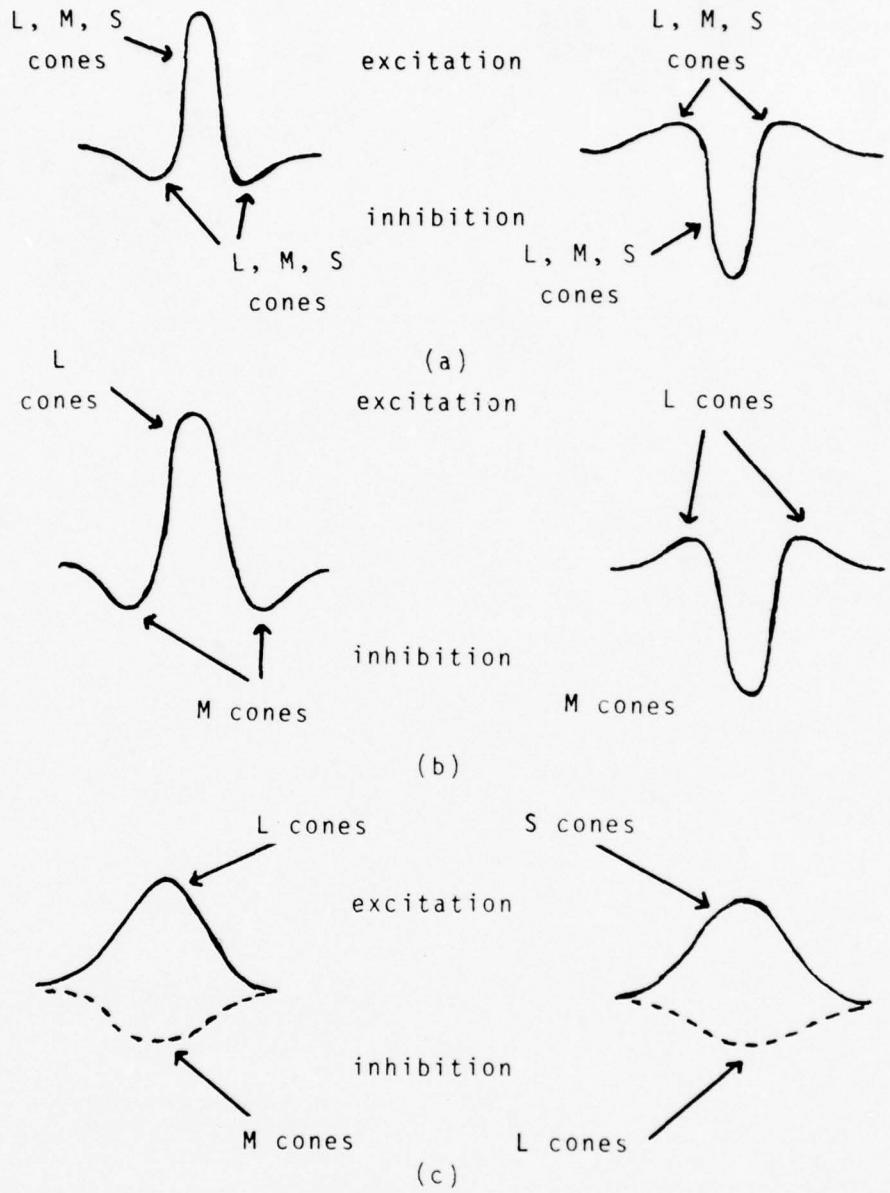



Fig. 1.3- Examples of the spatial organization of L G N cells.
a) Spectrally non-opponent cells.
b) First type of spectrally opponent cells.
c) Second type of spectrally opponent cells.

It is interesting to investigate whether this spatial organization can account for spatial effects. The color seen at one point in space at a particular time does not depend only upon the stimulus at that point in space and time but also on the stimuli at neighboring points. One very well known spatial interaction is called contrast. Simultaneous contrast refers to the fact that a surround of one color or brightness tends to induce the opposite color or brightness into a neighboring region. For example, in simultaneous brightness contrast, a grey appears black in a white surround and white in a black surround. In simultaneous color contrast, a grey appears reddish in a green surround, greenish in a red surround.

It is commonly supposed that the center-surround organization of non-opponent cells in the LGN accounts for simultaneous brightness contrast although this has been recently questioned [14]. On the other hand the spatial organization of opponent cells predicts the absence of pure color Mach Bands (induced by constant brightness stimuli) and does not account for simultaneous color contrast. Indeed, for pure color stimuli, the center and the surround act in the same direction and there is no center-surround antagonism to attenuate the responses to large stimuli. For example, in a (+R-G) cell in which the L-cones are excitatory in the center and the M-cones inhibitory in the surround, a green light will not evoke a large response but a red light of the same brightness will produce an increase in center excitation and a decrease in surround inhibition, thus evoking a much larger response. The center and the surround act in the same direction.

This is to be opposed to what happens in a non-spectrally opponent cell when switching from dark to light. We will say that it appears that simultaneous brightness contrast in part and simultaneous color contrast depend upon interactions that occur past the LGN [12].

We now turn to a precise description of the model that can be built on these experimental data.

# CHAPTER 2

# A HOMOMORPHIC MODEL FOR HUMAN COLOR VISION

## 2.1 Cone Stage

As described in chapter 1 there are three cone types involved in human photopic vision, each of them containing a pigment with different spectral absorption characteristics. From color-matching experiments, the CIE ("Commission Internationale de l'eclairage") derived a set of three color matching functions which, according to Grassman laws of color mixture, have to be a linear combination of the pigments absorption curves.

Of course there is an infinite number of such combinations but we are restricted by the fact that the resulting curves must be positive in the visible spectrum and have a single maximum at a wavelength close to the ones indicated by the methods referred to in chapter 1. Nonetheless, despite these important restrictions on the linear transformation there are quite a few systems of so-called "fundamentals" in the literature [20, 30, 61].

The system we chose is the one proposed by Stiles [47, 48]. The three resulting absorption curves $l(\lambda), m(\lambda), s(\lambda)$ are positive in the visible spectrum and have single maxima at respectively 575, 540 and 445 nm. They are shown in figure 2.1. The exact linear transformation that relates those spectral absorption curves to the CIE $x(\lambda)$, $y(\lambda)$ and $z(\lambda)$ curves is given in appendix C.

(a)

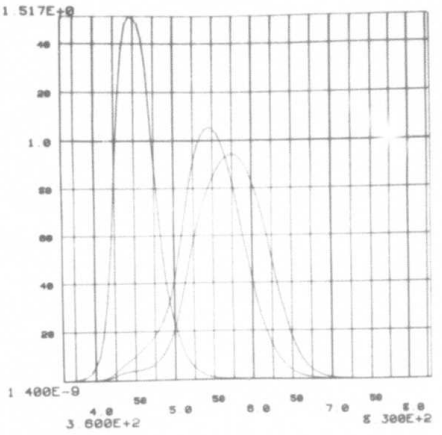



Fig. 2.1- The cones spectral absorption functions $l(\lambda)$, $m(\lambda)$, $s(\lambda)$.

(b)

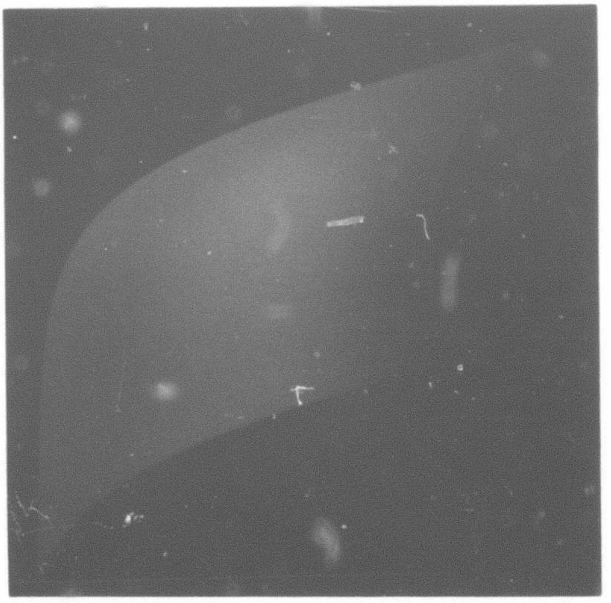


Fig. 2.2- A view of a constant A plane: the $C_1$ axis runs horizontally, the $C_2$ axis vertically. The curved boundaries are due to the logarithmic stage in the model and to the fact that only a limited gamut of colors can be reproduced in the primary system used.

We are now in position to give a precise description of the cone absorption stage of the model. Let $I(x,y,\lambda)$ be an image, where $x$ and $y$ are the spatial coordinates and $\lambda$ is the wavelength of the light. Then the absorption of this light energy by the three pigments yields the following three signals:

$$L(x,y) = \int I(x,y,\lambda) l(\lambda) d\lambda \qquad (2.1a)$$

$$M(x,y) = \int I(x,y,\lambda) m(\lambda) d\lambda \qquad (2.1b)$$

$$S(x,y) = \int I(x,y,\lambda) s(\lambda) d\lambda \qquad (2.1c)$$

where each integral is taken over the visible spectrum (from 380 to 800 nm).

The nonlinear response of the cones transforms those signals in the following way:

$$L^*(x,y) = \mathrm{Log}(L(x,y)) \qquad (2.2a)$$

$$M^*(x,y) = \mathrm{Log}(M(x,y)) \qquad (2.2b)$$

$$S^*(x,y) = \mathrm{Log}(S(x,y)) \qquad (2.2c)$$

It is interesting to mention that the problem of the exact type of the nonlinearity in human cones has yet to be solved. Cornsweet [6], for instance, advocates a nonlinearity of the type $U*I/(V+I)$ where U and V are constants and I is the input intensity. His reasons are based on a model of the chemical processes taking place during the cone absorption stage which predicts such a law. He also points out that over a large range of intensities this nonlinearity and the logarithm are in close agreement. Indeed it may be noted that the assumption of an exact logarithmic form for the nonlinearity is not a critical feature as far as psychophysics goes and that over large ranges of the variables other simple functions behave in much the same way. But, as we will see in paragraph 2.7,

from an image processing standpoint the situation is quite different and explains why we chose the logarithm apart from its computational simplicity.

Before we leave the topic of the absorption of light in the cone receptors and of the nonlinearity that immediatly follows it, we would like to mention one more fact. The light which is absorbed at the receptor level has already travelled through the optic media (cornea, aqueous humor, crystalline lens and vitreous humor) causing a blurred retinal image.This can be modelled as the result of a linear space-invariant filtering operation whose modulation transfer function has been measured [5] and is characterized by a high frequency attenuation. Another consequence is the introduction of chromatic aberrations where the image for shorter wavelengths lies slightly in front and for longer wavelengths slightly behind the retina. We will neglect this last effect in what follows.

## 2.2 Separation of chromatic and achromatic information

We showed in chapter 1 that there is strong neurophysiological evidence in favor of the existence of separate chromatic and achromatic channels in the human visual path, each gaining its information from the same receptors but processing it in different ways. The studies we referred to suggested a mathematical expression for the achromatic response, we will call it A, and the two chromatic responses, we will call them $C_1$ and $C_2$:

$$A = a(\alpha L^* + \beta M^* + \gamma S^*) = a(\alpha \mathrm{Log}(L) + \beta \mathrm{Log}(M) + \gamma \mathrm{Log}(S)) \qquad (2.3a)$$

$$C_1 = u_1(L^* - M^*) = u_1 \mathrm{Log}(L/M) \qquad (2.3b)$$

$$C_2 = u_2(L^* - S^*) = u_2 \mathrm{Log}(L/S) \qquad (2.3c)$$

A is a mathematical description of the response of (+Bl-Wh) and (+Wh-Bl) cells which add the Log-outputs of the L-, M- and S-cones. $C_1$ and $C_2$ are mathematical descriptions of the responses of (+R-G), (+G-R) and (+B-Y), (+Y-B) cells which subtract the Log-outputs of the L- and M-cones and L- and S-cones respectively. The relation between A, $C_1$, $C_2$ and L*, M*, S* is thus a linear one. The corresponding matrix 1/4 is given in appendix C. a, $\alpha$, $\beta$, $\gamma$, $u_1$, $u_2$ are constants which will be defined later.

## 2.3 Summation and contrast effects

Two important features of the way information is processed in the visual system have not yet been included in the model. The first one, which is responsible for the limited resolution of the human vision, is caused by the finite density of receptors in the retina and spatial summation of neural activity at different stages on the visual path (neural blur). As we mentioned earlier, another cause of this limited resolution is the optics of the eye (optical blur). The neural blur effect can be modelled as a linear system characterized by an attenuation of high frequencies [1, 5].

The second feature is the existence of lateral inhibition between receptors and neural cells at different places in the visual path (retina, LGN, cortex). This lateral inhibition effect by which a cell inhibits the activity of neighboring cells with a strength proportional to its excitation and decreasing with the distance is very much linear and, as shown by many authors [1, 8, 24], has a sharpening effect. It is characterized by a low-frequency

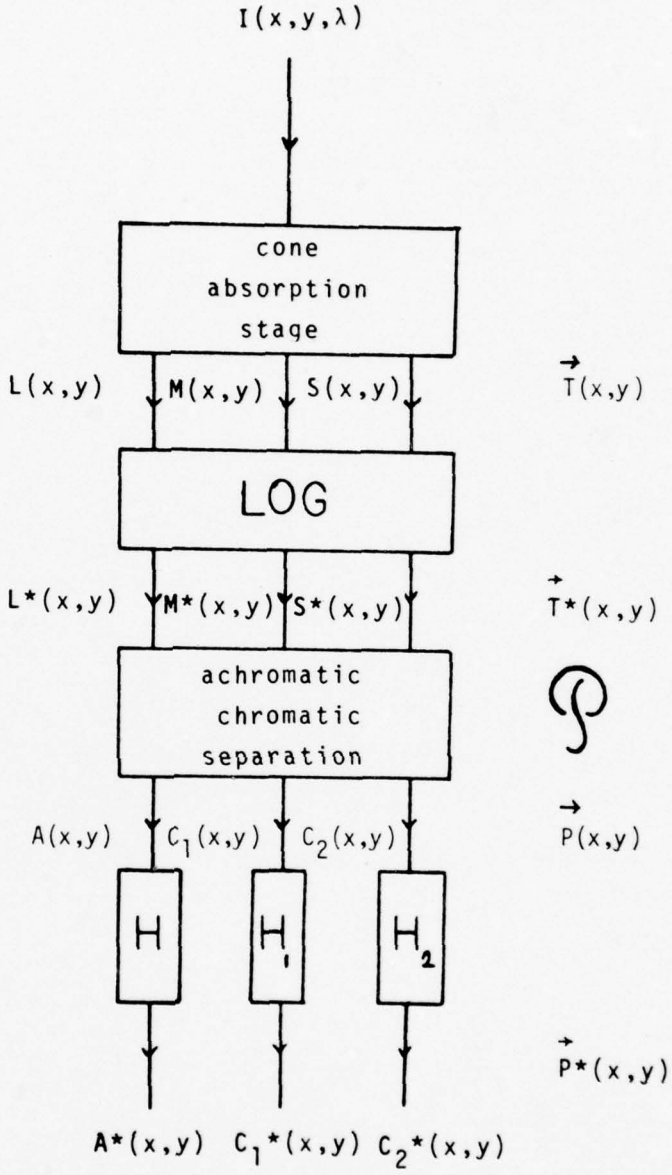



Fig. 2.3- The homomorphic model of color vision as suggested by neurophysiological and psychophysical evidence.

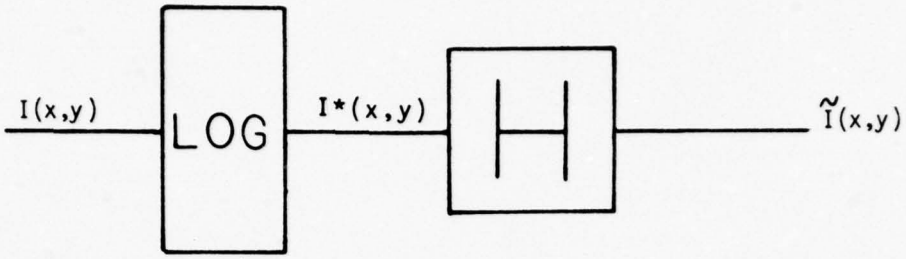



Fig. 2.4- Stockham's homomorphic model of achromatic vision.

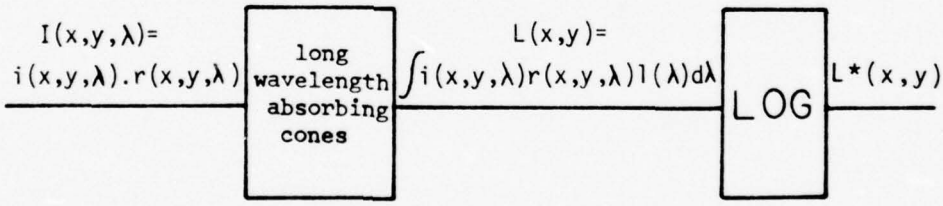



Fig. 2.5- The breakdown of reflectance-illumination separation. The model of color vision, just as the eye, does not exactly separate illumination from reflectance since the logarithmic state operates on integrals of products rather than on products.

attenuation.

Thus the total effect of the neural activity can be considered as the result of the cascade of those two systems, the neural blur attenuating high spatial frequencies and lateral inhibition attenuating low spatial frequencies.

Although both phenomena seem to be distributed all along the visual path and sometime difficult to separate from a physiological standpoint, we will assume here that we can approximate reality in a sufficiently precise manner by modelling those effects by three linear space invariant filters acting on signals A, $C_1$ and $C_2$.

Let $H(f_1,f_2)$, $H_1(f_1,f_2)$, $H_2(f_1,f_2)$ be the frequency responses of those filters where $f_1$ and $f_2$ are the two spatial frequencies. We will furthermore assume that each of these functions is the product of two functions corresponding to lateral inhibition and neural blur. In other words we dichotomize the two effects in every channel by considering each filter as the cascade of two filters modelling respectively lateral inhibition and neural blur. The corresponding point-spread functions are denoted by $h(x,y)$, $h_1(x,y)$ and $h_2(x,y)$. The perceptual interpretation of those filters will be discussed more in paragraph 6 and in chapter 3. They yield outputs $A^*$, $C_1^*$ and $C_2^*$ given by:

$$A^*(x,y)=\int\int A(x-\rho,y-\tau)h(\rho,\tau)\,d\rho d\tau \qquad (2.4a)$$

$$C_1^*(x,y)=\int\int C_1(x-\rho,y-\tau)h_1(\rho,\tau)\,d\rho d\tau \qquad (2.4b)$$

$$C_2^*(x,y)=\int\int C_2(x-\rho,y-\tau)h_2(\rho,\tau)\,d\rho d\tau \qquad (2.4c)$$

We are now in a position where we can draw a complete block diagram of our model. This is shown in figure 2.3.

## 2.4 Relation to other models

This model can be classified as an "opponent-color" model in the line of the theory proposed first by Hering [25]. He hypothesized three paired and opponent visual qualities, white-black, red-green and yellow-blue, independent of each other. His theory has had many followers who tried to make it quantitative in different ways. The most recent efforts in this direction have been made by Jameson and Hurvitch [26], Meessen [34], Shklover [46], Koenderink et. al. [31] and Frei [17]. But this is to our knowledge the first time that such a formalism is used for image processing and for an experimental study of color contrast effects.

Further this model is related to a model for achromatic vision that has been proposed and used for image processing by Stockham [50]. For an achromatic image the outputs of the cone absorption stage are by definition equal:

$$L(x,y)=M(x,y)=S(x,y)$$

and thus:

$$C_1(x,y)=C_2(x,y)=0$$

In that case only the achromatic channel A is active :

$$A=a(\alpha+\beta+\gamma)L^*(x,y)$$

and figure 2.3 reduces to figure 2.4 which is exactly Stockham's model. This implies a very important fact namely that we know the frequency response $H(f_1,f_2)$. It is the amplitude modulation transfer function measured by Davidson [8], Baudelaire [1] and others.

## 2.5 The problem of brightness perception

In order to discuss the perception of brightness, we will have to go back to the CIE observer and examine more closely another visual property he embodies and which bears on the perception of brightness. The property in question is expressed by the basic principle of photometry, the additive law for brightness also called Abney's law. According to this law, the condition for a match in brightness of two stimuli of radiances (spectral energy distributions) $R(\lambda)$ and $R'(\lambda)$ takes the form:

$$\int R(\lambda)V(\lambda)d\lambda = \int R'(\lambda)V(\lambda)d\lambda \qquad (2.5)$$

the integrals being taken as usual over the visible spectrum.

$V(\lambda)$ is a function of wavelength known as the relative luminous efficiency function. There are several methods for measuring this function, all involving monochromatic stimuli, many of them are reviewed in Guth [23]. They yield similar although not identical results.

The standardized $V(\lambda)$ curve, result of the CIE measurements, is everywhere positive in the visible spectrum and has a single maximum normalized to 1 at 555 nm.

Let us now investigate the meaning of equation (2.5) in terms of perception:

a) if two patches of light P and P' of radiances $R(\lambda)$ and $R'(\lambda)$ match in brightness then the two patches of radiances $cR(\lambda)$ and $cR'(\lambda)$ still match in brightness (scalability property).

b) if patch $P_0$ matches patch $P_1$ and also matches patch $P_2$ then the additive mixture of patches $P_1$ and $P_2$ will match patch $P_0$ increased to double its original radiance (additivity property).

For an observer who embodies this property and makes color-matches

according to Grassmann's laws of additive color mixture it can be shown that the function $V(\lambda)$ must be a linear combination of the color-matching functions and thus also of the cones spectral absorption curves. This fact together with the knowledge that the cone responses are nonlinear has lead people to assume that $V(\lambda)$ was one of the responses [17]. Of course this is in contradiction with our current knowledge about these fundamentals which we know to be peaking at 445, 540 and 575-580 nm while $V(\lambda)$ peaks at 555 nm.

Nonetheless our model should account for the measurements yielding to the function $V(\lambda)$. One method that can be used to measure the $V(\lambda)$ function is called the step by step method. Two juxtaposed monochromatic lights of slightly different wavelengths are viewed and the radiance of one is varied until the total "difference" between the patches is minimum. The wavelength difference can be made so small that the difference in color at this minimum setting is barely perceptible. This procedure is then repeated step by step along the spectrum.

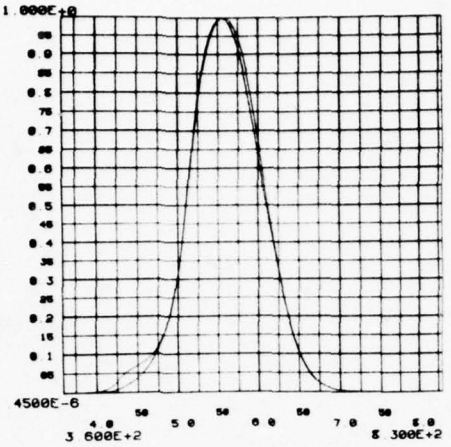
A relative luminous efficiency function can be derived from our model by a corresponding procedure as explained in appendix A. The resulting function that may be called $V^*(\lambda)$ is related to the cone spectral absorption curves by the following equation:

$$V^*(\lambda) = c\, l(\lambda)^{\alpha/(\alpha+\beta+\gamma)}\, m(\lambda)^{\beta/(\alpha+\beta+\gamma)}\, s(\lambda)^{\gamma/(\alpha+\beta+\gamma)} \qquad (2.6)$$

where the constant c has been adjusted to yield a maximum of 1. A comparison between $V^*(\lambda)$ and $V(\lambda)$ is shown in figures 2.6 and 2.7 on a linear and a logarithmic scale for the following values of the parameters $\alpha$, $\beta$ and $\gamma$:

$$\alpha = .612 \qquad \beta = .369 \qquad \gamma = .019$$

(a)




(b)

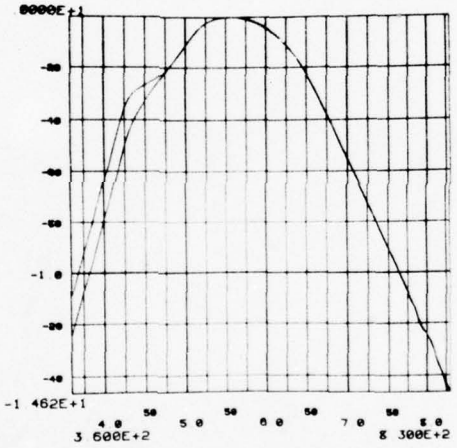



Fig. 2.6- Comparison between the C.I.E. relative luminous efficiency function $V(\lambda)$ and $V^*(\lambda)$ predicted by the model: on a linear scale (a) and a Log-scale (b). The values of $V^*(\lambda)$ are larger than the values of $V(\lambda)$ in the blue end of the spectrum, in qualitative agreement with results obtained by Judd.

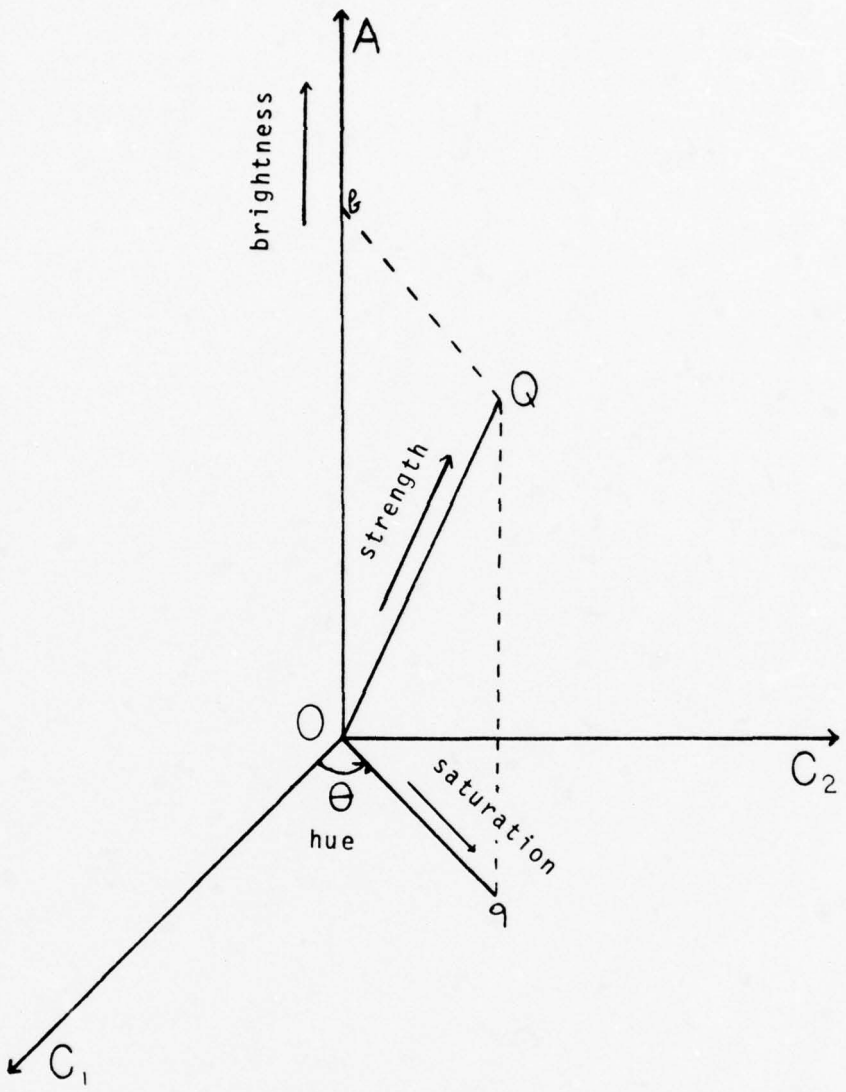



Fig. 2.7- The $(A,C_1,C_2)$ space as a three-dimensional vector space provides a quantitative definition of some perceptually important parameters. The origin O corresponds to a medium grey.

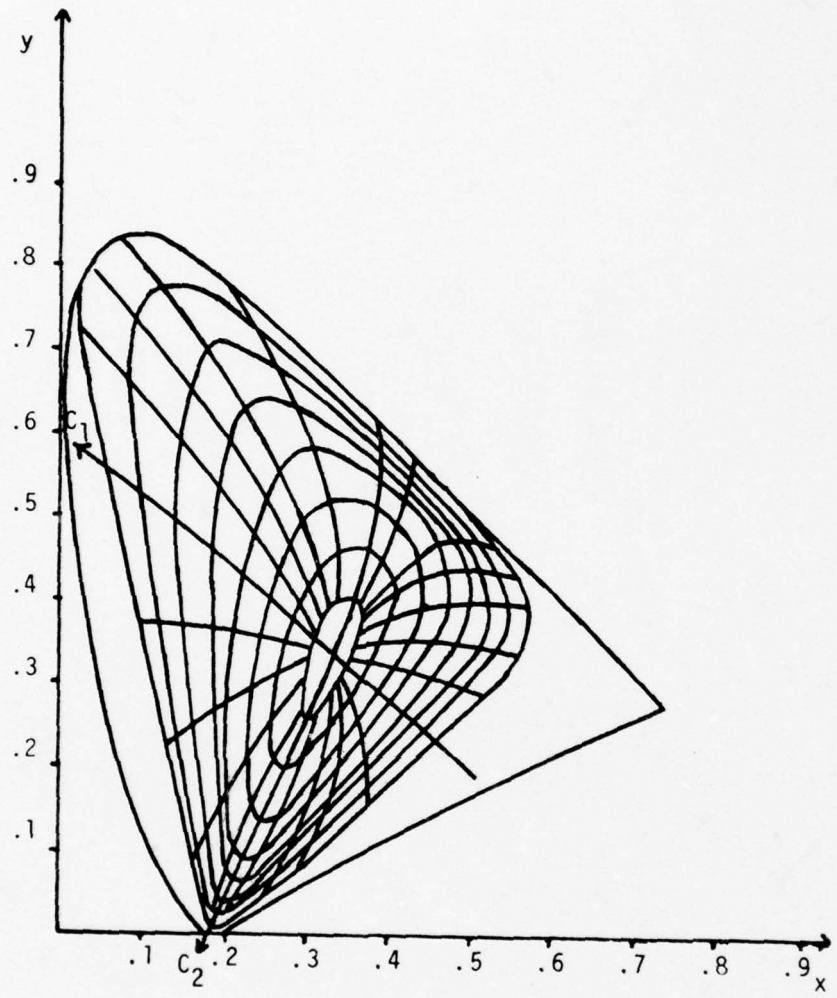



Fig. 2.8- Lines of constant hue and saturation predicted by the model in the C.I.E. chromaticity diagram.

(in particular $\alpha+\beta+\gamma=1$).

The reader will notice that the values of $V^*(\lambda)$ are greater than the values of $V(\lambda)$ in the blue end of the spectrum. This is not to be considered as a handicap for our model, on the contrary, since there is plenty of evidence that the CIE curve is underestimated in this part of the spectrum (see Wyszecki and Stiles [61], pp.429, 435).

Clearly we cannot say anything about Abney's law as long as we do not give a precise definition of brightness in the framework of our model. We showed that the visual information is conveyed from the retina to the brain via three independent channels one of them transmitting the achromatic or black and white information. This together with the aforementioned fact that for an achromatic image only this channel was excited makes it reasonable to say that it carries the brightness information. Thus, for a given color Q defined by its tristimulus values L, M, S, which in turn uniquely define the values of A, $C_1$, $C_2$, the brightness is proportional to A and hence to $\alpha Log(L)+\beta Log(M)+\gamma Log(S)$.

We can see at first that with this definition brightness obeys the scalability condition. Indeed if patch P of tristimulus values (L,M,S) matches in brightness patch P' of tristimulus values (L',M',S') then:

$$\alpha Log(L)+\beta Log(M)+\gamma Log(S) = \alpha Log(L')+\beta Log(M')+\gamma Log(S')$$

and patches of tristimulus values (cL,cM,cS) and (cL',cM',cS') still match in brightness.

On the other hand the additivity property is lost which means that in our model brightness perception does not completely follow

Abney's law. Now careful studies of the validity of the law have been conducted [23] and show that Abney's law does not hold especially for direct brightness matches. The authors further add that there does not seem to be an obvious replacement for the law because some colors are always subadditive (red and green) while others are subadditive at threshold and superadditive at high intensity levels (violet and green).

## 2.6 More about the perceptual interpretation of the $(A,C_1,C_2)$ space

As can be seen in figure 2.3, the model takes as its input an intensity image $I(x,y,\lambda)$ which is then processed to yield the three perceptual quantities $A(x,y)$, $C_1(x,y)$, $C_2(x,y)$ and eventually $A^*(x,y)$, $C_1^*(x,y)$ and $C_2^*(x,y)$ which we will conveniently describe as two three-dimensional vectors $\vec{P}(x,y)$ and $\vec{P}^*(x,y)$.

Let us ignore the spatial filtering which changes the values of the perceptual quantities but not their interpretation and let us also ignore for simplicity the spatial dependency. Every color Q of tristimulus values (L,M,S), which we will also describe as a three-dimensional vector $\vec{T}$, is also defined by a "perceptual vector" $\vec{P} = [A,C_1,C_2]^t$.

We have already discussed the interpretation of A with regard to brightness and we will just add to what has already been said that a value of zero corresponds to a neutral grey while increasing positive values of A correspond to brighter and brighter stimuli and decreasing negative values correspond to darker and darker stimuli.

If we now consider colors for which the value of A is constant one may ask how is the variation of $C_1$ and $C_2$ related to the

perception of hue and saturation. We will say that the perception of saturation is related to the distance bQ of the point Q to the A-axis while the perception of hue is related to the angle $\theta$ (figure 2.7).

When we make these assumptions we are defining a euclidian vector space structure on the $(A,C_1,C_2)$ space which allows us to talk about adding vectors and about distances and angles. We will show in chapter 5 that the norm does not have to be euclidian and that other choices are possible. We will come back to the interpretation of vector addition with regard to the structure of images in paragraph 2.7 and concentrate now on what may be called the metric.

Saturation is proportional to the distance Oq (see figure 2.7) and is thus related to the sum of the activities in the $C_1$ and $C_2$ channels while hue is related to the ratio of these activities. The distance OQ of the point representing the color to the origin O of the $(A,C_1,C_2)$ space is also of interest:

$$OQ = (Oq^2 + Ob^2)^{1/2}$$

It is a measure of both the brightness and the saturation of the color or equivalently of the total activity in all three channels. Following Stockham we might call it the strength of the color Q.

At constant brightness, lines of constant hue are straight lines radiating from and perpendicular to the A-axis and lines of constant saturation are circles centered on this axis. It is interesting to represent those lines in the CIE (x,y) plane as in figure 2.8. The resulting network has the same general shape as the one obtained from color-order systems such as Munsell or Din. To

give an idea of what we mean by planes of constant A or constant brightness, such a plane is shown in figure 2.2. Each point of coordinates $(C_1,C_2)$ in this plane is represented by a perceptual vector $\vec{P} = [A,C_1,C_2]^t$ where A is the same for all points. This perceptual vector uniquely defines a tristimulus vector $\vec{T} = [L,M,S]^t$ which is represented on a film by different dyes concentrations. The curved boundaries are due to the limited gamut of colors reproducible by the display primaries.

Another very important property of the $(C_1,C_2)$ plane is the following: it is a close approximation to a uniform chromaticity plane which has the characteristic that at a given point, the locus of all points corresponding to a just noticeable difference in hue is a circle of radius 1 centered on the point. Similarly, the $(A,C_1,C_2)$ space is a uniform perceptual space in the sense that at a given a point, the locus of all points corresponding to a just noticeable difference in perception (brightness or chromaticity) is a sphere of radius 1 centered at that point. This is achieved as explained in appendix B by adjusting the values of parameters a, $u_1$ and $u_2$ of equations 2.3a, 2.3b and 2.3c, respectively.

## 2.7 Discussion of the label homomorphic

Homomorphic is an adjective that is used in mathematics where people study special functions called homomorphisms. Grossly speaking they are functions from one set to another that are compatible with the laws of composition of elements in both sets. For example the logarithm function is a homomorphism of the multiplicative group of real positive numbers onto the additive group of all real numbers.

In the engineering field homomorphic systems that use these abstract ideas have been studied by Oppenheim and others [35] and proven to be very powerful tools by which the basic theory of linear systems has been extended.

In particular, Stockham [50, 51] proposed a multiplicative homomorphic model for processing achromatic images that matched both the structure of images and of the human visual system. The multiplicative process of image formation (product of illumination $i(x,y)$ and reflectance $r(x,y)$) is harmoniously undone by the Log sensitivity of the retinal receptors that maps this product into a sum of two components that can later on be separated by linear filtering techniques since in general their spatial frequency content tends to be distinct.

This analysis applies to black and white images but partly breaks down for color since it ignores image wavelength content. Indeed, the image formation is a multiplicative process that yields the product of the scene illumination $i(x,y,\lambda)$ and reflectance $r(x,y,\lambda)$ but, before the visual system processes this product with a logarithm, the light is linearly absorbed by the three cone pigments. Hence no longer do we have the logarithm of a product but the logarithm of a sum of products which is considerably different. This situation is depicted in figure 2.5 for the L-cones.

At first glance it seems that the color model is not able to account for the automatic gain control properties of the human visual system as the achromatic model is. Nonetheless, let us show that the model of figure 2.3 does account for the ability of the visual system to discard the illuminant. Indeed suppose we

uniformly increase the absolute level of illumination by letting $i(x,y,\lambda)$ become $ki(x,y,\lambda)$ where k is a constant larger than 1. Then it is clear that $L(x,y)$, $M(x,y)$ and $S(x,y)$ are changed to $kL(x,y)$, $kM(x,y)$ and $kS(x,y)$ respectively. In other words all three cone outputs increase by the same amount and thus the chromaticity channels outputs $C_1$ and $C_2$ don't change while the brightness channel output A is biased by the constant $a\mathrm{Log}(k)$ (equations 2.3a to 2.3c), which in turn is readily eliminated by lateral inhibition (or synonymously spatial filtering with low frequency attenuation).

Now suppose we not only change the absolute level of illumination but also the color of the illuminant so that we can approximately write that $L(x,y)$, $M(x,y)$ and $S(x,y)$ become respectively $lL(x,y)$, $mM(x,y)$ and $sS(x,y)$ where l,m and s are different constants. In other words the cone outputs are changed by different amounts uniformly over the field.

*The brightness channel output is biased by the constant* $a(\alpha\mathrm{Log}(l)+\beta\mathrm{Log}(m)+\gamma\mathrm{Log}(s))$, again readily eliminated by lateral inhibition. But, in contrast to the first case, the chromaticity channels outputs are biased respectively by $u_1\mathrm{Log}(l/m)$ and $u_2\mathrm{Log}(l/s)$ which are non zero. In order that this bias be readily eliminated we must assume that lateral inhibition also occurs in those channels. Chapter 3 will be devoted to show that this is indeed the case.

As a summary, when we go from achromatic vision to color vision we lose the exact match between image structure as a product of two components and model or processor structure. Our model, like the eye, does not exactly separate reflectance from illumination.

Nonetheless we have been able to show that it can explain the tremendous adaptation ability of the human visual system and chapters 4 and 5 will show that it can be used to perform useful processing tasks on images.

From this standpoint, let us model the tristimulus vector $\vec{T}(x,y) = [L(x,y),M(x,y),S(x,y)]^t$ as a product of two components $\vec{t}(x,y) = [l(x,y),m(x,y),s(x,y)]^t$ and $\vec{t}'(x,y) = [l'(x,y),m'(x,y),s'(x,y)]^t$ so that one has:

$$\vec{T}(x,y) = \vec{t}(x,y).\vec{t}'(x,y)$$

or equivalently:

$$L(x,y) = l(x,y)l'(x,y)$$

$$M(x,y) = m(x,y)m'(x,y)$$

$$S(x,y) = s(x,y)s'(x,y)$$

It is clear from figure 2.3 that if $\vec{P}(x,y)$ and $\vec{P}'(x,y)$ are the perceptual vectors corresponding to $\vec{t}(x,y)$ and $\vec{t}'(x,y)$ respectively then $\vec{T}(x,y) = \vec{t}(x,y).\vec{t}'(x,y)$ corresponds to $\vec{P}(x,y) + \vec{P}'(x,y)$. Thus, we now have the interpretation of vector addition in the $(A,C_1,C_2)$ space, it simply corresponds to multiplication of tristimulus values.

Let us see how this remark coupled with our assumption of independant spatial filtering on the perceptual quantities $(A,C_1,C_2)$ is of interest to an image processor. Suppose that $\vec{t}(x,y)$ is slowly varying and $\vec{t}'(x,y)$ rapidly varying, then the same will be true of $\vec{P}(x,y)$ and $\vec{P}'(x,y)$ respectively. But since they are now additively combined, it will be possible to separate them, at least approximately, by means of linear filtering thus opening the door to many applications, some of them described in chapter 4.

We now realize how the notion of the model as a 3-dimensional homomorphic system is important. It accounts for the automatic gain control properties of the eye and introduces new meaningful color image processing techniques.

# CHAPTER 3

# A STUDY OF SPATIAL EFFECTS IN THE CHROMATICITY CHANNELS

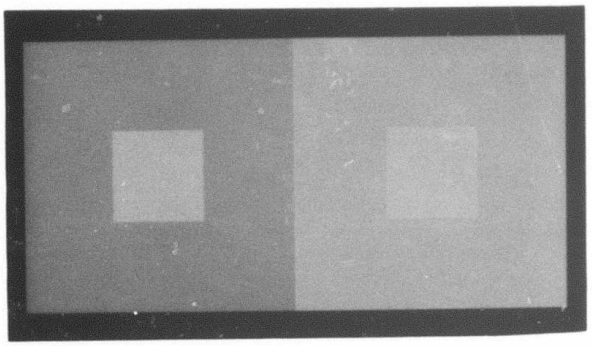
## 3.1 Visual adaptation and contrast effects

There are two types of visual adaptation. One sort happens with sudden changes in the intensity or chromaticity of the illuminant and is a fairly slow process (several minutes) related to the bleaching of visual pigments [6]. The other sort that shows the ability of the human visual system to adjust instantaneously to the illuminant is related to neural interactions past the retina. This is in this last type of adaptation we are interested here and will attempt to model.

The phenomenon of brightness constancy is such a type of adaptation and shows the tremendous ability of the visual system to discard the spatial variations in intensity of the illuminant. It has been shown by several authors [6, 50] that a Log type sensitivity at the receptor level followed by a spatial linear filtering characterized by an attenuation of low frequencies would account for this ability. Experimental studies conducted by several authors [1, 8] have confirmed this hypothesis by actually measuring the modulation transfer function of the neural network.

It should be noted that this low frequency attenuation not only accounts for brightness constancy but also for brightness contrast

(a)



(b)

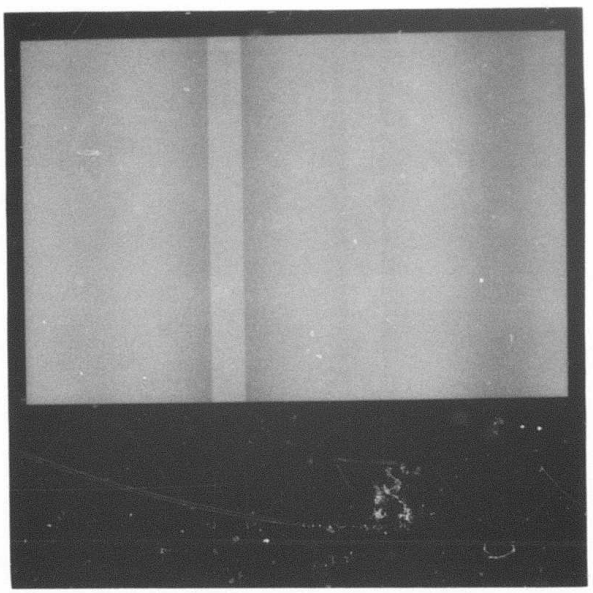
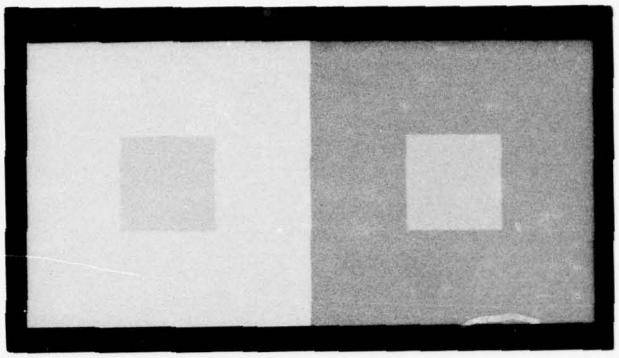


Fig. 3.1- Simultaneous color contrast in the $C_1$ -channel.

a) The two small squares have the same tristimulus values.

b) The two vertical stripes have the same tristimulus values.

(a)



(b)

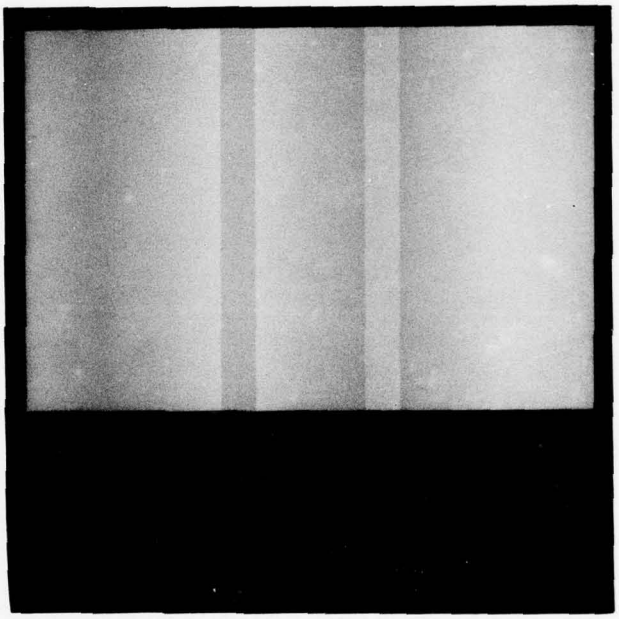


Fig. 3.2- Simultaneous color contrast in the $C_2$-channel.

a) The two small squares have the same tristimulus values.

b) The two vertical stripes have the same tristimulus values.

effects where the brightness at one point is influenced by the brightness of neighboring points and does not correspond to what might be expected from the distribution of light intensities if there was no neural interaction. Mach Bands, simultaneous brightness contrast illusions, Herman Grids, are example of such effects and the interested reader is referred to Ratliff [37], Baudelaire [1] and Baxter [2] for a more complete coverage.

Not only is the visual system able to discard variations in intensity of the illuminant but also variations in color. The light of an ordinary light bulb is very yellow and the light outside on a clear sunny day is very blue but in both cases we will call a piece of white paper white even though the spectral content of the light it reflects to our eye is very different in both cases. This phenomenon is called color constancy.

Just as color constancy corresponds to brightness constancy, brightness contrast has its conterpart with color contrast. Here it is the hue and saturation at one point which is influenced by hues and saturation at neighboring points (figures 3.1 and 3.2). Of course both effects (brightness and color contrast) can occur simultaneously and it might be difficult to separate them.

Land's retinex theory is an attempt to account for brightness and color constancy [32].In this theory the log-outputs of neighboring receptors are pair-wise subtracted in an independent manner for every cone system. By adding these data along a path connecting two points on the retina the ratio of the corresponding cone outputs can be computed. This permits a comparison independently of a uniform or quasi-uniform illumination. But the

result does not depend on comparisons elsewhere and therefore this theory does not account for either brightness or color contrast effects. Another drawback of this theory is its lack of algebraic structure which makes it difficult, if not impossible, to use to make quantitative predictions.

Empirical formulations, such as the von Kries coefficients, have been presented to account for the perceptual appearance of colors and color induction [29, 38, 52] but it seems that the homomorphic approach which already models brightness effects fairly well is a more powerful and general tool.

Before getting more into these questions, one might ask what are the practical implications of gaining an understanding of how colors interact and how the eye adapts to the illuminant.

One possible answer can be found among the many people producing colored products. They have to worry about how their final product is going to look and, in order to avoid bad surprises such as colors that do not look the way they are supposed to because another color has been placed beside, those people have had to study color induction problems empirically. A good quantitative model for such effects would obviously be welcome (Judd 1940, [29]).

Another answer can be found by looking toward the film industry. Color films which have a fixed spectral response cannot adapt to different illuminants and if a film balanced for, for example, tungsten lamp light is used to take an outdoor scene the result will be an objectionable color cast on the final print which can only be avoided by trial and errors in the photo lab during the printing process. In that case again a good understanding of the

adapting ability of the human eye would be welcome and a step forward toward an automatization of such a process as well as a solid basis for the handling of digital color images.

Since a spatial filtering on the chromaticity signals in our model characterized by an attenuation of low spatial frequencies would account for color constancy and color contrast, an experiment was designed to test this hypothesis.

## 3.2 The ideas involved in the design

The validity of the multiplicative homomorphic model of achromatic vision has been questioned for patterns with sharp edges. Davidson and Whiteside [9] indicated that edge effects were incompatible with the model predictions. Their results have been confirmed by Baudelaire who concluded that the existence of a highly nonlinear edge oriented mechanism centrally located (that is to say past the LGN) and influencing the whole visual field was most likely. The neural interaction network responsible for most brightness effects on the other hand is quite linear in the absence of edges.

We will also adopt here a careful attitude toward edges and use smooth patterns when measuring $H_1(f_1, f_2)$ and $H_2(f_1, f_2)$. In other words, we will only use patterns containing a few low spatial frequencies. Also we assume for simplicity that the lateral inhibition postulated in the two chromaticity channels of our model is the same in all directions which means that the frequency responses $H_1$ and $H_2$, and the corresponding point-spread functions $h_1$ and $h_2$, are circularly symmetric. Consequently, they are entirely

defined by their cross-section which allows us to experiment only with one-dimensional patterns.

These patterns were displayed on the face of a high quality color television monitor, part of a Comtal display unit, driven by a general purpose digital computer. The display was carefully calibrated so that it was easy to relate precisely the digital intensities to the intensities of the red, green and blue lights emitted by the three phosphors on the face of the tube. The calibration procedure is explained in appendix D.

If $S_i(x)$ $(i=1,2,3)$ is the spatial modulation of intensity for the ith phosphor and $p_i(\lambda)$ the spectral distribution of the light emitted by this phosphor, the image generated on the face of the tube is

$$I(x,\lambda) = \sum_{i=1}^{3} S_i(x)\, p_i(\lambda)$$

The corresponding tristimulus values are

$$L(x) = \sum_{i=1}^{3} S_i(x) \int p_i(\lambda)\, I(\lambda)\, d\lambda$$

and two other similar equations for $M(x)$ and $S(x)$. If we denote by $L_i$, $M_i$ and $S_i$ $(i=1,2,3)$ the tristimulus values of the light emitted by the ith phosphor (for example $L_i = \int p_i(\lambda)\, I(\lambda)\, d\lambda$ ) then this can be conveniently rewritten as a vector equation

$$\vec{T}(x) = U_i \ \vec{S}(x)$$

where $\vec{S}(x) = [S_1(x), S_2(x), S_3(x)]^t$, $\vec{T}(x) = [L(x), M(x), S(x)]^t$ and $U_i$ is a three by three matrix whose column vectors are the tristimulus values of the phosphors. $U_i$ is computed as explained by the following diagram:

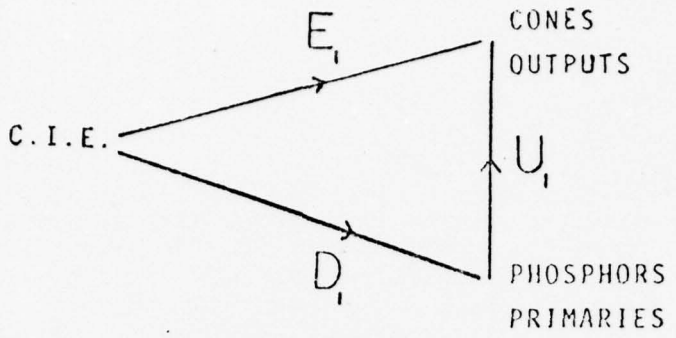



Fig. 3.3- Computation of the $U_i$ matrix from matrices $E_i$ and $D_i$.

$E_i$ is the matrix given in appendix C that relates the cone spectral sensitivities to the CIE color matching functions $x(\lambda)$, $y(\lambda)$ and $z(\lambda)$. $D_i$ has for column vectors the chromaticity coordinates of the lights emitted by the phosphors and is thus characteristic of those phosphors and given by the builder. Thus $E_i$ depends only on the model (which cone sensitivities were chosen) and $D_i$ depends only on the phosphors. The only relationship between the two is through the white.

Indeed $D_i$ and $E_i$ are not completely defined as long as a reference white has not been chosen. For the monitor it corresponds to equal drive signals on the three guns (or equivalently to equal digital intensities) and for the eye model to equal tristimulus values L, M, S. In other words if $\vec{W}$ is the three-dimensional vector whose components are the chromaticity values (x,y,z) of the chosen white (for the television monitor it is usually D6500) then one has

$$\vec{I} = E_i \ \vec{W} \qquad \text{and} \qquad \vec{I} = D_i \ \vec{W}$$

so that

$$\vec{I} = U_i \ \vec{I}$$

where $\vec{I}$ is the vector $[1,1,1]^t$.

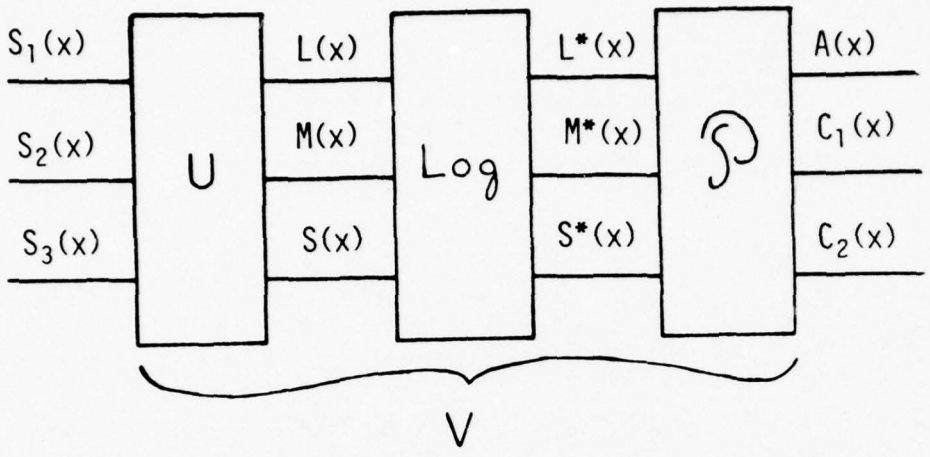



Fig. 3.4- The system V transforms the three dititat images $S_1$, $S_2$, $S_3$ into the perceptual signals A, $C_1$, $C_2$.

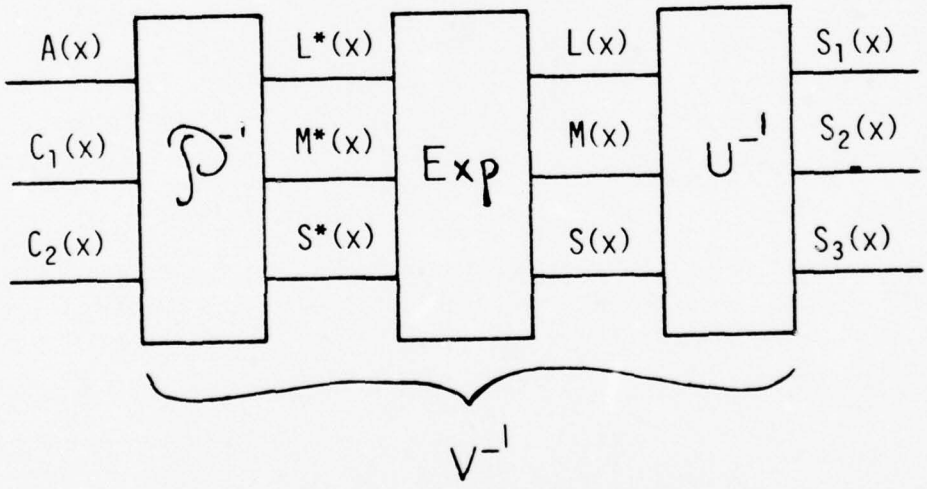



Fig. 3.5- The system $V^{-1}$ transforms the three perceptual signals A, $C_1$, $C_2$ back into the three digital images $S_1$, $S_2$, $S_3$.

The simple algebraic procedure used is explained in appendix C.

This completely defines $U_1$ and thus the relationship between the three digital images $\vec{S}(x)$ and the tristimulus vector $\vec{T}(x)$.

It should be pointed out that we have neglected in this computation the effect of the eye optics which as we recall is equivalent to a linear low pass filter. But since we are experimenting with smooth patterns our approximation is justified [1].

What we want to achieve with the images $\vec{S}(x)$ is to generate perceptual signals $A(x)$, $C_1(x)$ and $C_2(x)$ in the visual system of a human observer that allow us to study the frequency responses $H_1$ and $H_2$. This situation is summarized in figure 3.4.

In other words the problem is, knowing what A, $C_1$ and $C_2$ have to be, what are $S_1$, $S_2$ and $S_3$. This question is easily answered by inverting the model of figure 3.4. If we condense all the steps or subsystems between $(S_1,S_2,S_3)$ and $(A,C_1,C_2)$ and represent them as a system V then, since it is a cascade of invertible subsystems, V itself is invertible to $V^{-1}$ as pictured in figure 3.5. So if we know what A, $C_1$, $C_2$ signals we want it is easy to define the required $S_1$, $S_2$, $S_3$ signals by applying $V^{-1}$ to them.

The $(A,C_1,C_2)$ pattern we used are given by:

$$[A=a \quad C_1=c_1+k(\sin(2\pi fx)+\alpha\sin(6\pi fx)) \quad C_2=c_2] \tag{3.1a}$$

$$[A=a \quad C_1=c_1 \quad C_2=c_2+k(\sin(2\pi fx)+\alpha\sin(6\pi fx))] \tag{3.1b}$$

where a, $c_1$, $c_2$, k, $\alpha$ are constants, f is the spatial frequency and x the spatial coordinate. The frequency content of such patterns is limited to the 0-frequency (DC) component for the two channels that are not probed (the achromatic channel is always one of those) and

to frequencies f and 3f as well for the chromatic channel that is tested.

As we mentioned earlier those patterns were generated in a general purpose digital computer that performed the required simulation of the inverse of V and were displayed on the face of a high quality television monitor driven by the computer.

The examples of those patterns shown in this thesis use as the final recording media a piece of photographic film. Unfortunately there is no good quantitative model which, given the spectral distribution of the input light and the characteristics of the film and of the printing process used, can predict the resulting tristimulus values experienced by a human observer when looking at the final print illuminated by a given light. One attempt has been made in that direction by Wallis [57] but his results, although encouraging, show that more research effort needs to be put in this domain.

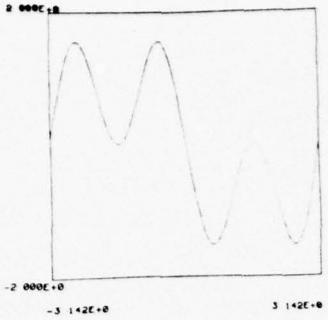
Consequently in opposition to the case of the television monitor it is very difficult to relate accurately digital intensities to tristimulus values for the film media and the pictures in this thesis showing color illusions are intended only to give an idea of the actual stimuli seen on the face of the color television monitor.

The shape of the curve of equation

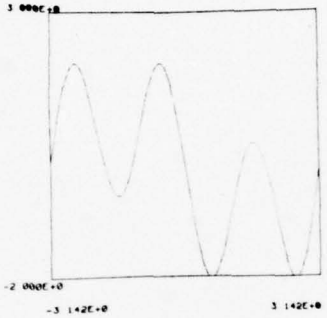
$$C(x)=c+k(\sin(2\pi fx)+\alpha\sin(6\pi fx)) \qquad (3.2)$$

is shown in figure 3.6 for $\alpha=1$, $\alpha>1$ and $\alpha<1$. Setting $\alpha$ equal to 1 in equation (3.2) allows us to test for low spatial frequency attenuation in

(a)



(b)



(c)

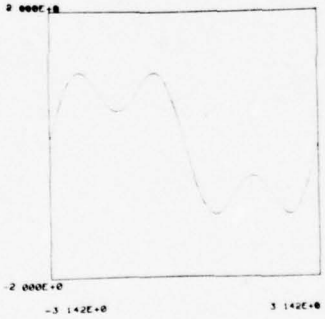


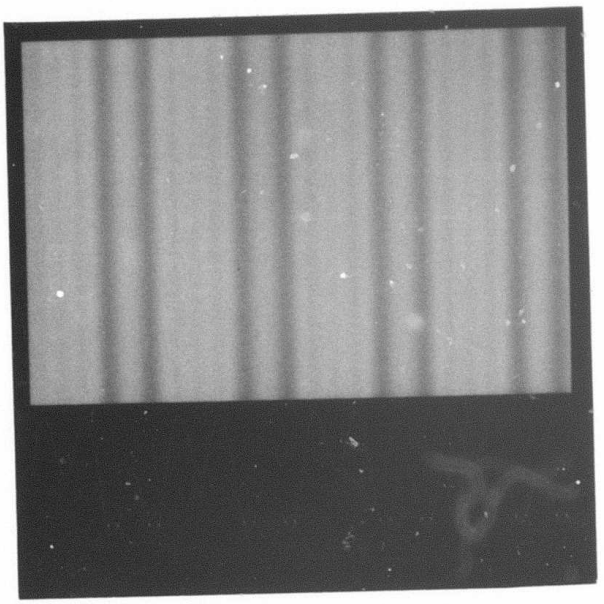
Fig. 3.6- The shape of the curve of equation $C(x)=c+k(\sin(2\pi fx)+\alpha\sin(6\pi fx))$:

a) $\alpha=1$, the leftmost trough and the rightmost peak are at the same level.

b) $\alpha>1$, the leftmost trough is at a lower level than the rightmost peak.

c) $\alpha<1$, the leftmost trough is at a higher level than the rightmost peak.

(a)



(b)

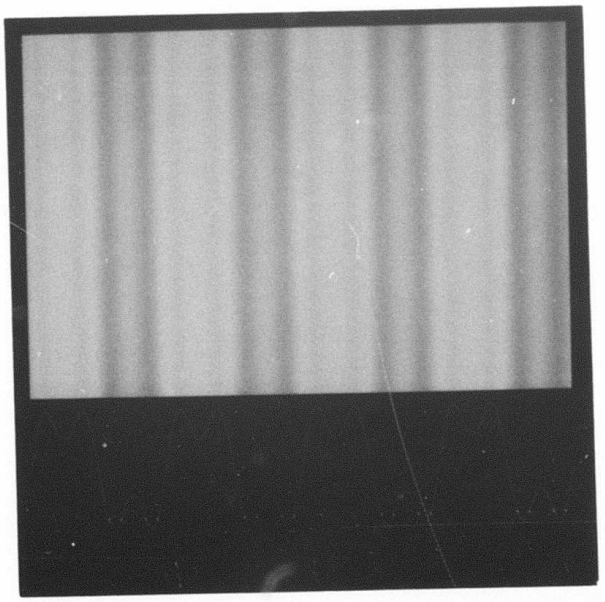
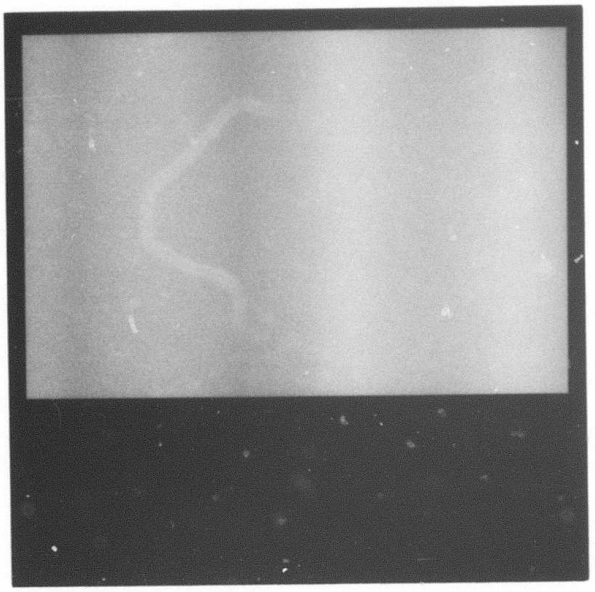


Fig. 3.7- Test pattern corresponding to a sinusoidal stimulus in the $C_1$ -channel and constant stimuli in the A- and $C_2$ -channels.
a) The value of $\alpha$ in equation (3.1a) is 1 and a simultaneous color contrast illusion can be observed.
b) The value of $\alpha$ in equation (3.1a) is 0.5 which corresponds approximately to the cancellation of the illusion.

(a)



(b)

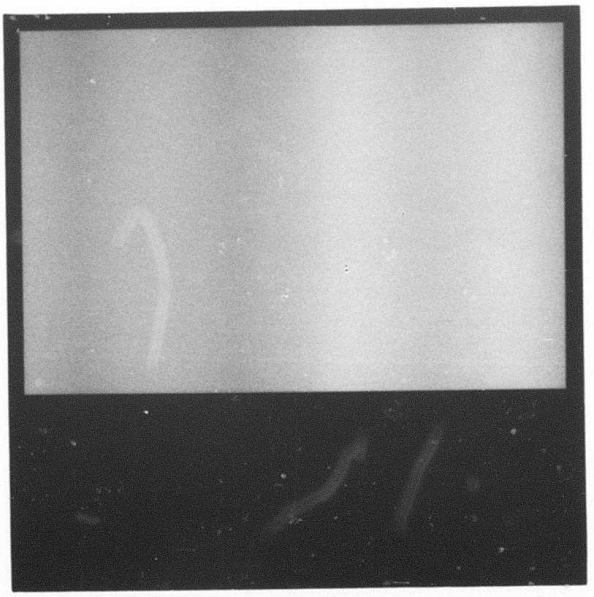


Fig. 3.8- Test pattern corresponding to a sinusoidal stimulus in the $C_2$ -channel and constant stimuli in the A- and $C_1$ -channels.
a) The value of $\alpha$ in equation (3.1b) is 1 and a simultaneous color contrast illusion can be observed.
b) The value of $\alpha$ in equation (3.1b) is 0.6 which corresponds approximately to the cancellation of the illusion.

the channel of interest. Indeed if we assume linear spatial filtering on C(x) characterized by an amplitude modulation transfer function K(f) the actual stimulus experienced is

$$C^*(x) = K(0)c + k(K(f)\sin(2\pi fx) + \alpha K(3f)\sin(6\pi fx))$$

which can be rewritten as

$$C^*(x) = K(0)c + kK(f)(\sin(2\pi fx) + \alpha(K(3f)/K(f))\sin(6\pi fx)) \qquad (3.3)$$

so if K(3f) is greater than K(f) and $\alpha$ equals 1 we should experience an illusion as indicated by figures 3.6, 3.7 and 3.8.

This illusion can actually be cancelled by setting $\alpha$ to K(f)/K(3f) (equation (3.3)). This is the key to the measurement of the low frequency attenuation in the chromaticity channels.

For a fixed spatial frequency f, patterns corresponding to values of $\alpha$ between 1 and .1 were presented to a set of observers and the value of $\alpha$ which cancelled the color illusion experienced when $\alpha$ was equal to 1 was a measure of the ratio K(f)/K(3f). The experience was repeated for four spatial frequencies ($f_1$=0.142, $f_2$=0.284, $f_3$=0.568, $f_4$=1.136 cycles/degree) on both chromaticity channels. The experiment is described in more details in appendix E.

It should be emphasized that all this reasoning is based upon the assumption of linearity of the lateral inhibition and that this experiment is also a test of that property.

## 3.3 Results of the experiment

We found that in the range of frequencies studied the following relations held:

$$H_1(3f) = 2.0H_1(f) \qquad (3.4a)$$

$$H_2(3f) = 1.7H_2(f) \qquad (3.4b)$$

which should be compared with Baudelaire's result for H:

$$H(3f) = 2.0H(f) \qquad (3.4c)$$

and Davidson's result:

$$H(3f) = 2.2H(f) \qquad (3.4d)$$

These relations imply a power-law for H, $H_1$ and $H_2$ in the frequency band tested and thus a linear relationship on a Log-Log scale. Thus our conclusion is that there is a low frequency attenuation in the chromaticity channels and the relative amplitude modulation of frequencies f and 3f is about the same in the red-green channel as in the achromatic and a little less in the yellow-blue channel for the band of frequencies studied. This is in agreement with other studies [21] which concluded that the contrast sensitivity of the yellow-blue mechanism is less than the contrast sensitivity of the red-green and brightness mechanisms. Of course relations (3.4a) and (3.4b) do not completely determine $H_1$ and $H_2$ and in particular say nothing about very low frequencies (below .1 cycle/degree) and about frequencies higher than about 4 cycles/degree.

But, just as for brightness effects the important band of frequencies is from about .1 to 6 cycles/degree [1], for color effects the important band is from about .1 to 2-4 cycles/degree.

Nonetheless we were able to get at least a rough idea of the shapes of $H_1$ and $H_2$ at higher frequencies as follows. First of all in the experiment on the $C_2$-channel (yellow-blue) at frequency $f_4$=1.136 cycle/degree four out the five observers reported seeing the illusion in reverse that is to say reported seeing the stripes in the middle of the blue bars bluer than the stripes in the middle

of the yellow bars for $\alpha$ set to 1 while the fifth observer reported seeing the illusion much weaker than for the other three frequencies. This effect is called simultaneous color similitude [12]. In our case, this means that for the majority of observers $H_2(3f_4)$ was less than $H_2(f_4)$ (see figure 3.6) thus allowing us to roughly estimate the position of the maximum at 2 cycles/degree. The same phenomenon also happened in the study of the $C_4$-channel but for a higher frequency (2.272 cycles/degree) which also allowed us to roughly position the maximum of $H_4$ at about 4 cycles/degree.

Past those maxima there seems to be a rolloff toward higher frequencies. We indeed know that the neural network between the retina and the visual cortex is characterized by two important properties: lateral inhibition which applies only over a limited range and whose transfer function is characterized by a low-frequency attenuation and is flat for high frequencies (high-pass filter) is the first. Neural summation that occurs at different levels and whose equivalent transfer function is characterized by a high frequency attenuation (low-pass filter) is the second. Although it is clear that those two processes are physiologically combined their effects on visual patterns are quite different and it seems justified to separate them conceptually.

We mentioned earlier that the contrast sensitivity of the visual neural system had been measured by contrast threshold methods by Campbell and Green [5]. Baudelaire used their data for frequencies higher than 8 cycles/degree as a good estimate of the modulation transfer function of the neural summation for the brightness channel. We will also use them here in the following

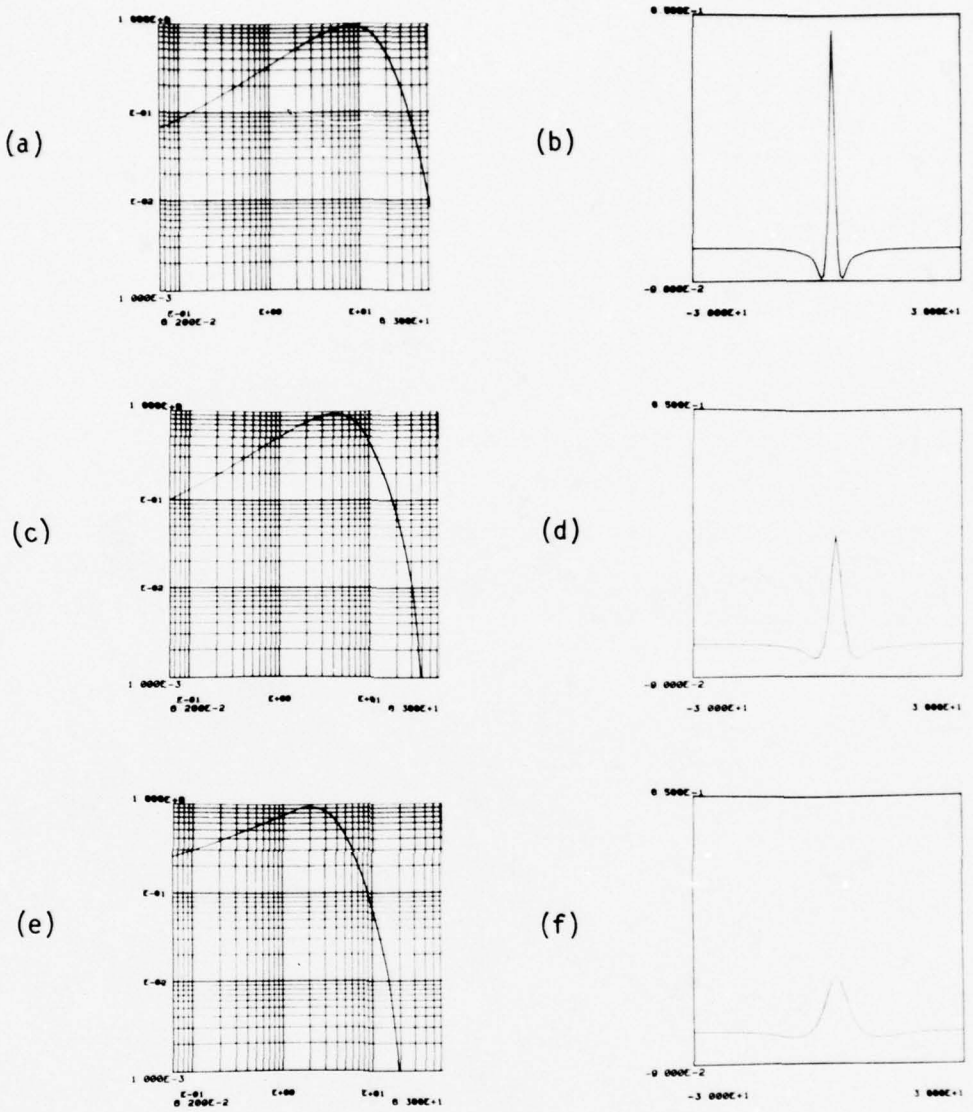



Fig. 3.9- The frequency responses H(f), $H_1(f)$ and $H_2(f)$ of the filters operating on the A, $C_1$ and $C_2$ signals (a, c, e), respectively, and (b, d, f) the corresponding point-spread functions h(x), $h_1(x)$ and $h_2(x)$. The frequency responses are plotted on a log-log scale, the point-spread functions for a span of one degree of visual angle.

way.

Let K(f) be the Campbell and Green data transformed as follows. K(f) is flat and equal to 1 from 0 to 8 cycles/degree and for higher frequencies is the previously mentioned data. This is what we will take, after Baudelaire, as the modulation transfer function equivalent to the neural summation for the brightness channel.

We will assume that $K_1(f)$, the modulation transfer function equivalent to neural summation in the $C_1$-chromaticity channel, is a shifted version of K(f) ($K_1(f)=K(2f)$). This is in agreement with our finding of a maximum at about 4 cycles/degree for $H_1(f)$ and with the fact that visual acuity is poorer for pure colors than for brightness.

In the same way we will assume that $K_2(f)$ is also a shifted version of K(f) ($K_2(f)=K(4f)$) which, again, is in agreement with our finding of a maximum at about 2 cycles/degree for $H_2(f)$ and with the fact that visual acuity is poorer for yellow-blue colors than for red-green colors [56].

Next we will combine our experimental data on the low-frequency part of $H_1$ and $H_2$ with the modified Campbell and Green data. The resulting frequency responses are shown in figure 3.9 as well as the corresponding point spread functions.

## 3.4 Relation to the color Mach Bands problem

The question of whether it is possible to induce Mach Bands with chromatic gradients only (keeping brightness constant) is still controversial. Some people reported negative results [15, 18, 55] tending to show that lateral inhibition was not present in the

chromaticity channels thus leaving the well verified fact of simultaneous color contrast completely unexplained. Some other people [22, 28] reported that colored Mach Bands were visible on color gradients with constant luminance but that those effects were weaker than in the achromatic case.

Let us see how our findings might provide us with an explanation for this controversy. Comparing equations (3.1a) to (3.1c) and taking into account the fact that the peak frequencies are 8,4 and 2 cycles/degree for H, $H_1$ and $H_2$ respectively we can approximately write

$$H_1(f) = H(2f)$$

$$H_2(f) = H(4f)$$

which state that the two chromatic frequency responses are shifted versions of the achromatic frequency response. In the space domain these equations are equivalent to

$$h_1(x) = 1/2h(x/2) \qquad (3.5a)$$

$$h_2(x) = 1/4h(x/4) \qquad (3.5b)$$

Qualitatively, as can be seen in figures 3.9b, 3.9d and 3.9f , the effect is to broaden the point-spread functions $h_1$ and $h_2$ by a factor of 2 and 4 respectively with respect to h as well as scaling them down by the same factors.

Coming back to the colored Mach Bands problem, relations (3.5a) and (3.5b) suggest that they should be broader than the achromatic Mach Bands and consequently less visible.

# CHAPTER 4

## APPLICATION OF THE MODEL TO IMAGE PROCESSING

### 4.1 Scanning color images in and out of a computer

In order to test the model on real life color images it was necessary first to digitize such images and second to use a scanning-out procedure where a piece of color film was exposed and processed in a fixed and controlled way. This was done to minimize display equipment and film nonlinearities as well as the effects of the printing process on the results of the experiments and make sure that what we were looking at were actual consequences of the digital processing and not of film processing.

A color film (Kodak Ektacolor Professional Film, Type S) was exposed outdoors yielding a color negative after development. We made sure that the dynamic range of the scene fitted that of the film so that the only nonlinearity in this process was represented by the gammas of the D-LogE curves of the three dyes present in the negative after development. They were measured by exposing the same day a calibrated grey-scale and plotting the three curves showing transmission density on the negative versus reflection densities on the grey-scale. The next step in the scanning-in process was to prepare three black and white prints from this color negative. They were obtained by exposing panchromatic paper (Kodak Resisto Rapid Pan) from the color negative through three Wratten filters (Red 92,

Green 59 and Blue 47B). It turned out that when doing so we had to use the whole dynamic range of the paper and that consequently we were not using only the linear part of its D-LogE curve as in the case of the negative but also the parts usually referred to as the "toe" and the "shoulder" thus introducing another nonlinearity. Those three prints were then scanned independently on a reflectance scanner and the two nonlinearities mentioned earlier compensated for by software.

The net result of this set of operations was three digital images corresponding to the red, green and blue light intensities in the original scene. Since we compensated for all nonlinearities in the scanning-in process, the numbers in the computer were actually proportional to the light intensities incident on the film when the shutter was opened.

The scanning-out process was done in two ways. The first, intended to provide a "quick look" at the results of an experiment was simply to display, the red, green and blue images on the face of the television monitor we used for the psychophysical experiments described in chapter 3. The second, intended to get a high-quality hard-copy, consisted in displaying the red, green and blue images on the face of a very high quality cathode ray tube (CRT) and exposing sequentially a piece of color film (Kodak Vericolor S Film) through three filters (Wratten filters number 25, 58 and 47B). The film was then taken to the photo lab where it was developed according to a fixed procedure described in appendix D.

The results of the total scanning-in and out procedures are

(a)

Fig. 4.1- The 512 by 512 resolution original color image "BECKY".

(b)

Fig. 4.2- The 256 by 256 resolution original color image "CAR-PORT".

shown for two real life color images "BECKY" and "CAR-PORT" in figures 4.1 and 4.2.

## 4.2 Practical implications

Let us see now how the model developed in the first three chapters can be used to process those images. The obvious problem is that the input to the model (figure 2.3) is $I(x,y,\lambda)$, a function of wavelength as well as of spatial coordinates, while the only data that are available to us are three images $R(x,y)$, $G(x,y)$ and $B(x,y)$, results of the scanning-in process, where the wavelength information has been lost.

But we notice that we already solved this problem in chapter 3. Indeed in the case of the television monitor we knew the chromaticity coordinates of the lights emitted by the phosphors and thus the matrix transformation from the tristimulus values corresponding to those primaries into cone responses. In the case of the CRT we know the characteristic of the light emitted by its phosphors (CIE C illuminant) and this entirely defines the chromaticity coordinates of the light coming through the three Wratten filters used. Thus in this case again we can easily compute the matrix $U_2$ such that

$$[L(x,y),M(x,y),S(x,y)]^t = U_2\ [R(x,y),G(x,y),B(x,y)]^t$$

We can now redraw figure 2.3 as figure 4.3. $U_2$ is given in appendix C.

Since we have a way of mapping color images defined by $\vec{I}(x,y)=[R(x,y),G(x,y),B(x,y)]^t$ into the perceptual space as $\vec{P}(x,y)=[A(x,y),C_1(x,y),C_2(c,y)]^t$ and that furthermore this mapping

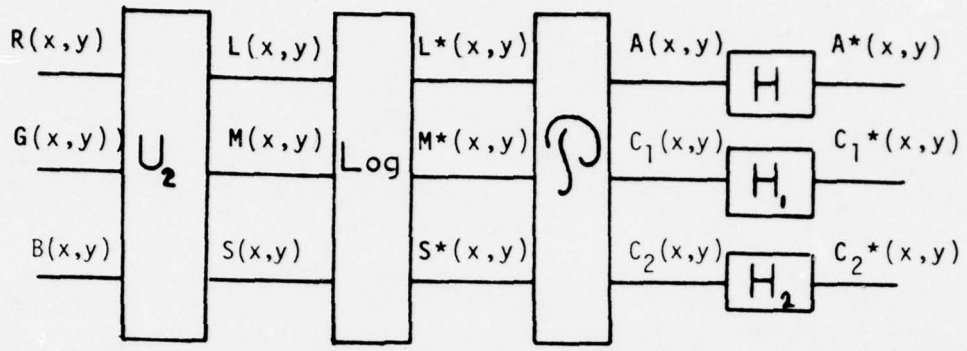



Fig. 4.3- The homomorphic model of color vision when the input image has been analysed in some set of primaries.

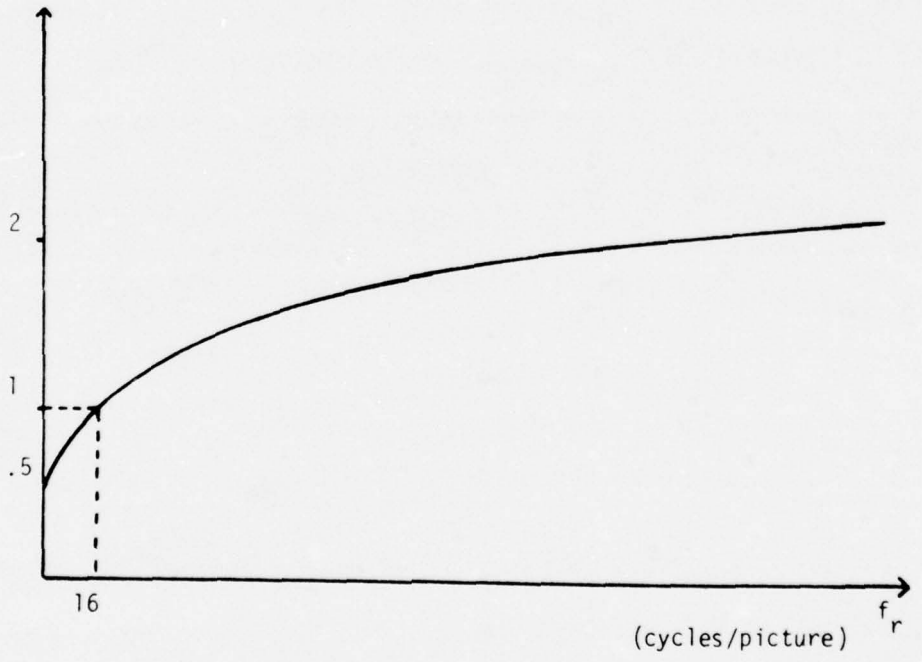



Fig. 4.4- Cross-section of the frequency response of the filter used to enhance brightness information.

Fig. 4.5- "BECKY" processed to yield a constant A (constant brightness) image.

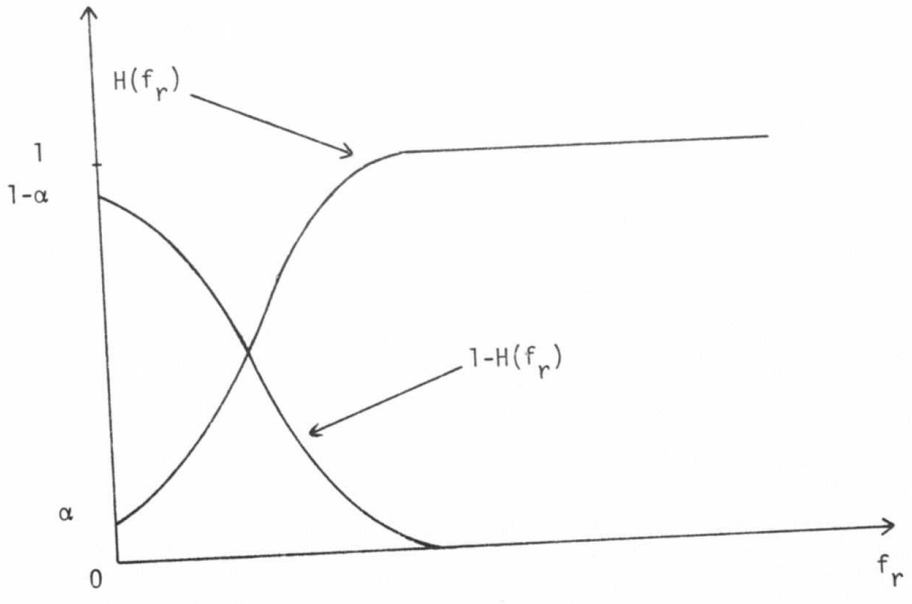



Fig. 4.6- Illustration of the unsharp masking idea.

is invertible, it is tempting to process $\vec{P}(x,y)$ because we know that by doing so we will actually be processing the quantities that are perceptually important for a human observer. This idea is summarized in figure 4.7. From equations 2.3a to 2.3c it is easy to derive

$$\hat{L}^* = L^* + (\hat{A} - A)/a + (\beta/u_1)(\hat{C}_1 - C_1) + (\gamma/u_2)(\hat{C}_2 - C_2) \quad (4.1a)$$

$$\hat{M}^* = M^* + (\hat{A} - A)/a + ((\beta - 1)/u_1)(\hat{C}_1 - C_1) + (\gamma/u_2)(\hat{C}_2 - C_2) \quad (4.1b)$$

$$\hat{S}^* = S^* + (\hat{A} - A)/a + (\beta/u_1)(\hat{C}_1 - C_1) + ((\gamma - 1)/u_2)(\hat{C}_2 - C_2) \quad (4.1c)$$

4.3 Brightness processing

Using those ideas we first experimented with the brightness information. For instance figure 4.5 shows the effect of setting $A(x,y)$ to a constant across the whole image which now looks like a cartoon with very little variation in brightness but where variations in color have been preserved. The reader may argue for example that the sky or the horse head look brighter than the rest of the scene. He is however confusing brightness with strength (see figure 2.7). From equations 4.1a to 4.1c it is easy to understand what happens to the tristimulus values. Since we have

$$\hat{A} = \text{constant} = c \qquad \hat{C}_1 = C_1 \qquad \hat{C}_2 = C_2$$

they can be rewritten as

$$\hat{L}^* = L^* + A'$$

$$\hat{M}^* = M^* + A'$$

$$\hat{S}^* = S^* + A'$$

where $A' = (c - A)/a$ and thus

$$\hat{L} = L \exp(A')$$

$$\hat{M} = M \exp(A')$$

$$\hat{S} = S \exp(A')$$

or in vector notation

$$\vec{\hat{T}} = \exp(A')\ \vec{T} \qquad (4.2)$$

If we remember that these quantities are functions of the spatial coordinates x and y, the interpretation of our processing is as follows. At a point (x,y) on the image where A is close to c, the chosen constant brightness, A' is small and thus exp(A') is close to 1 and we do not change the tristimulus vector $\vec{T}$ very much. If A is less than c, then A' is positive, exp(A') greater than 1, and we increase the tristimulus values at that point. Finally if A is greater than c, A' is negative, exp(A') less than 1, and we decrease the tristimulus values at that point.

In the same line of ideas we can also perform linear, space-invariant filtering on A(x,y) where the frequency response H of the filter is circularly symmetric and has a *low frequency* attenuation as shown in figure 4.6. Equation (4.2) can be rewritten as

$$\vec{\hat{T}} = (1/\exp(-A'))\ \vec{T} \qquad (4.3)$$

Noticing that $-A' = (A - \hat{A})/a$ and that, as shown in figure 4.6, $A - \hat{A}$ is a low-pass filtered version of A, equation (4.3) expresses nothing else than what is known to photographers as unsharp masking. Of course there is nothing that can hamper us from choosing a different frequency response, for example with not only a low frequency attenuation but also a high frequency amplification thus obtaining a processing much more general than unsharp masking. This processing, in the example just described, can be understood from equation (4.3) but also from the standpoint of our concluding

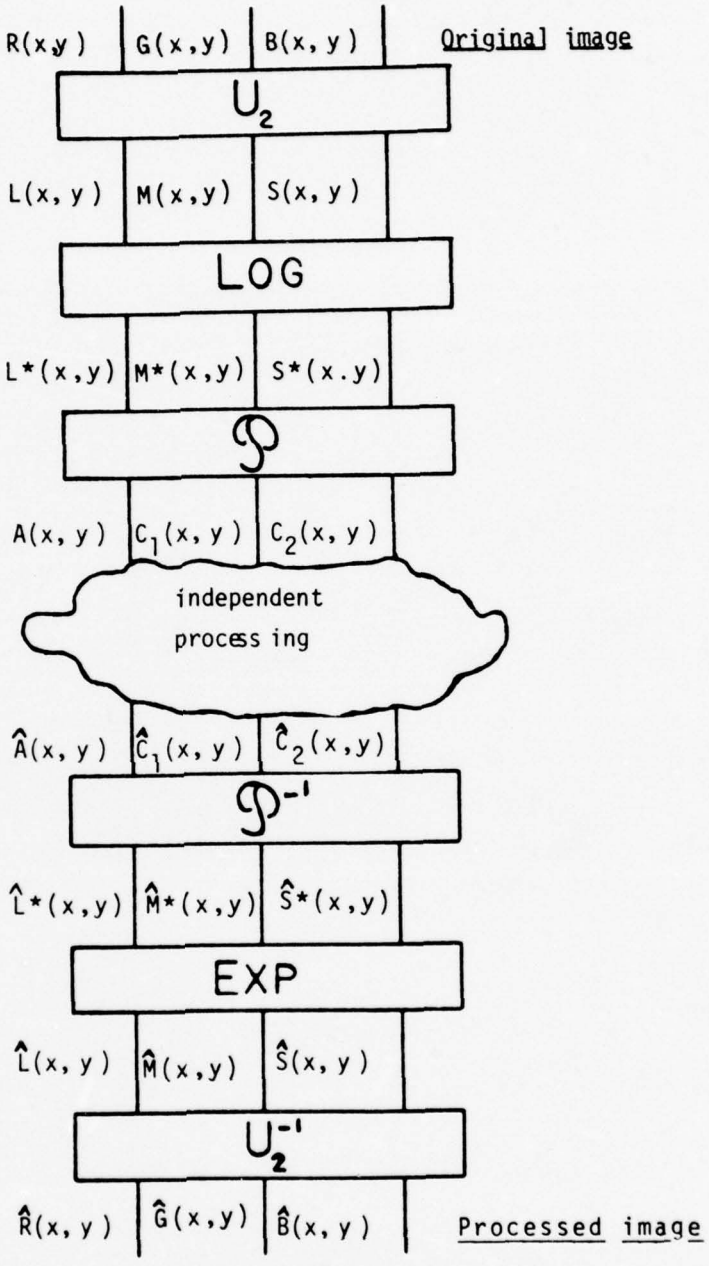



Fig. 4.7- The processing is based on a mapping of images into a "perceptual" space defined by the model where one operates on perceptually meaningful quantities. After processing, the image is mapped back into the system of display primaries.

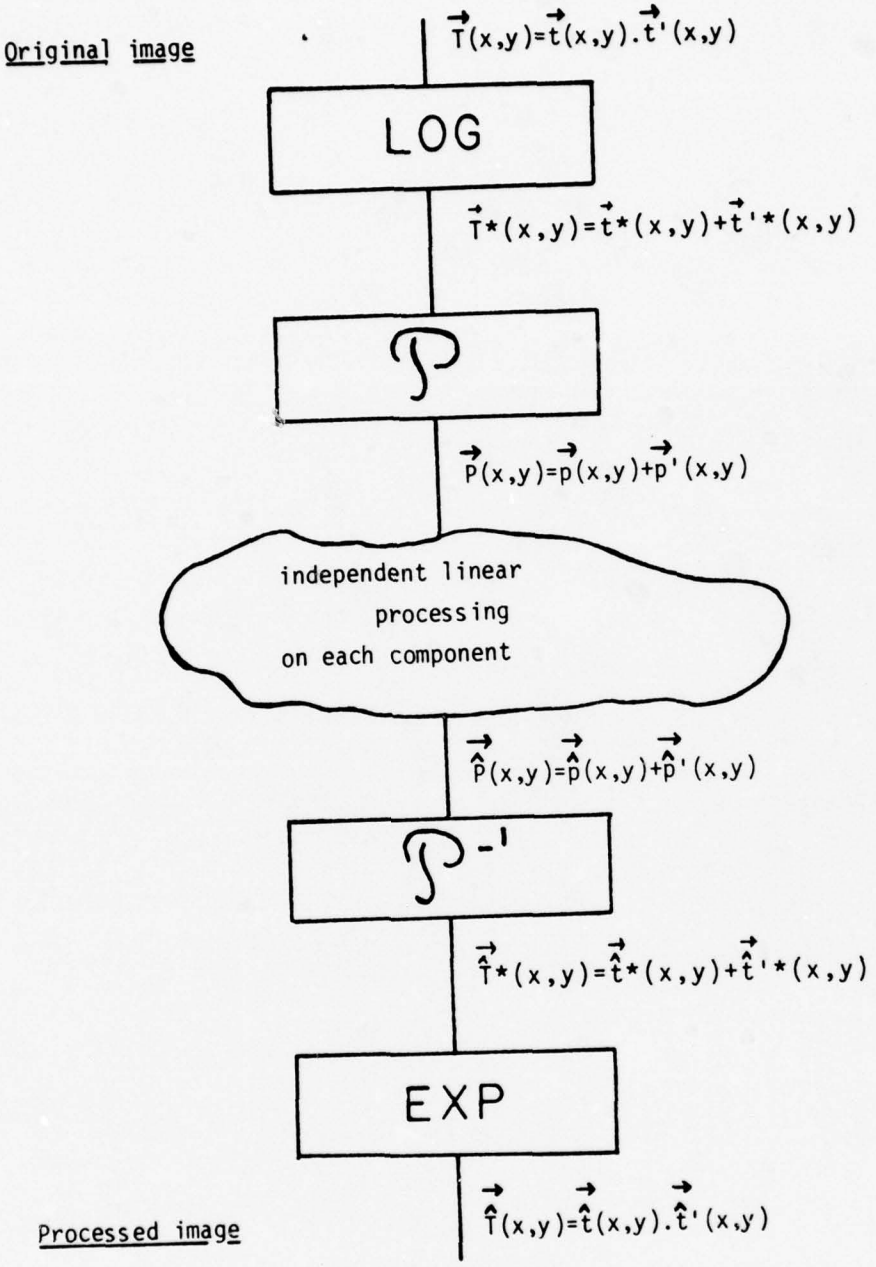


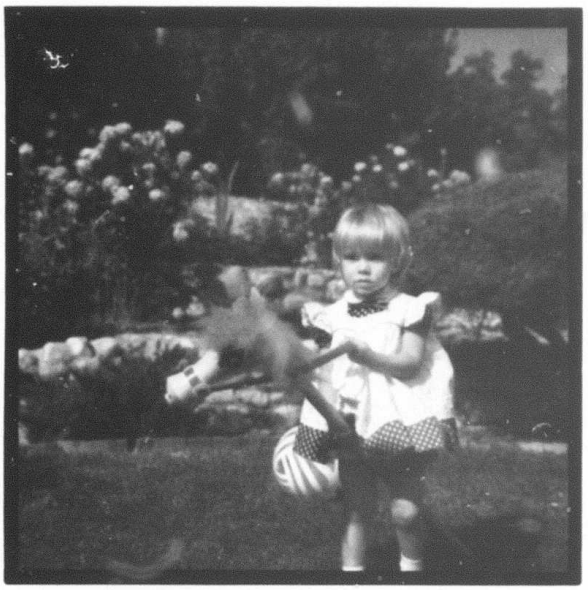

Fig. 4.8- Homomorphic interpretation when linear processing is performed in the "perceptual" space. Notice at every step the harmony between the processing and the structure of the processed quantities.

(a)

(b)

Fig. 4.9- Brightness enhancement experiment on (a) "BECKY" and (b) "CAR-PORT".

(a)



(b)

Fig. 4.10- Chromatic enhancement experiment on (a) "BECKY" and (b) "CAR-PORT".

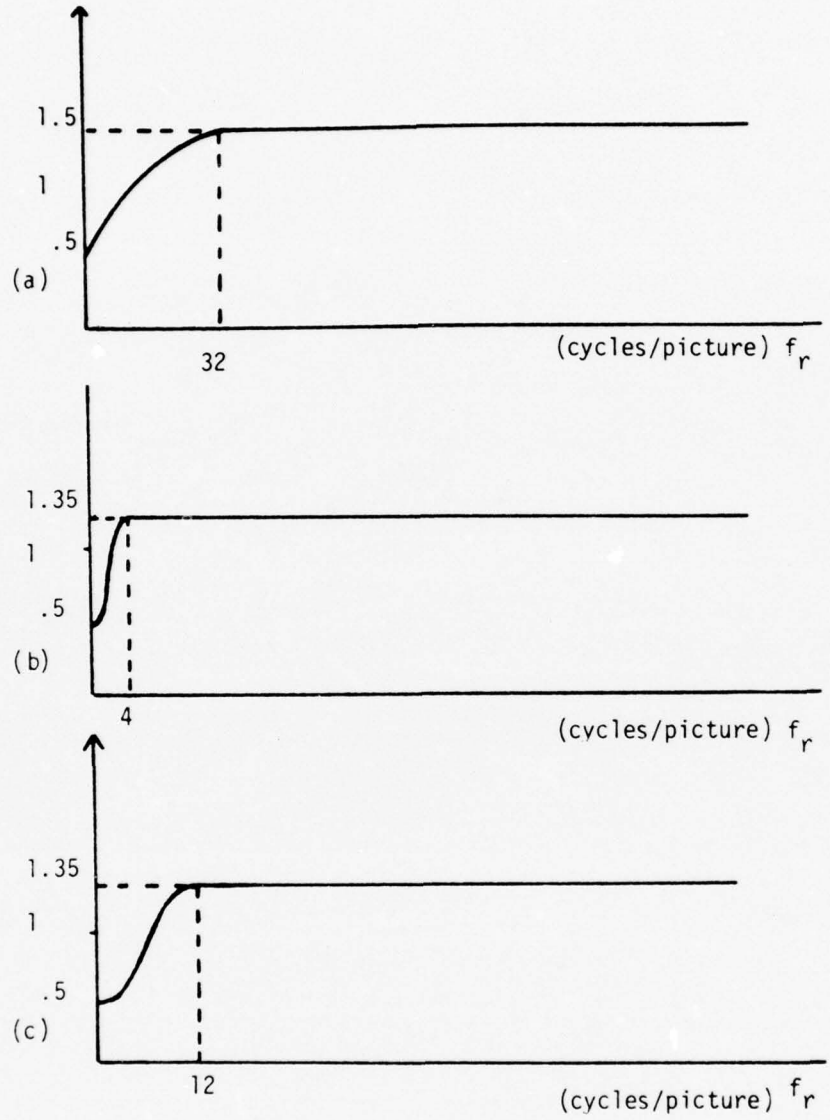
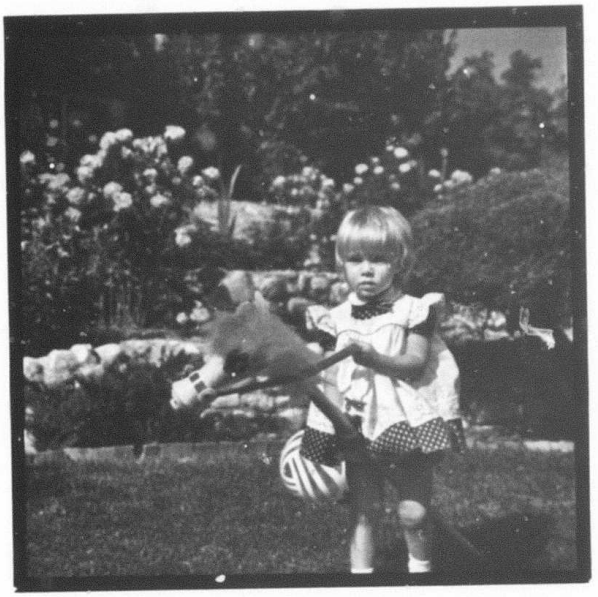



Fig. 4.11- Cross-sections of the frequency responses of the filters used to enhance the chromatic information in (a) the $C_1$- and $C_2$- channels for "CAR-PORT", in (b) the $C_1$- channel for "BECKY" and in (c) the $C_2$- channel for "BECKY".

(a)



(b)

Fig. 4.12- Brightness and chromatic enhancement experiment on (a) "BECKY" and (b) "CAR-PORT".

remarks in chapter 2.

Indeed if we model $\vec{T}=[L,M,S]^t$ as a product of a slowly varying component $\vec{t}=[l,m,s]^t$ and a rapidly varying component $\vec{t}'=[l',m',s']^t$, figure 4.7 can be redrawn as figure 4.8. $\vec{T}(x,y)$, $\vec{t}(x,y)$ and $\vec{t}'(x,y)$ are mapped through the model respectively onto $\vec{P}(x,y)=[A(x,y),C_1(x,y),C_2(x,y)]^t$, $\overline{p}(x,y)=[b(x,y),c_1(x,y),c_2(x,y)]^t$ and $p'(x,y)=[b'(x,y),c_1'(x,y),c_2'(x,y)]^t$ such that $\vec{P}(x,y)=p(x,y)+p'(x,y)$. In this particular case we do not process the chromatic signals and thus $\hat{C}_1 = c_1$, $\hat{C}_2 = c_2$, $\hat{C}_1' = c_1'$ and $\hat{C}_2' = c_2'$. The brightness signal A(x,y) is the sum of a slowly and rapidly varying components b(x,y) and b'(x,y) which are processed differently by our filter. The slowly varying component b, equal to $a(\alpha l^*+\beta m^*+\gamma s^*)$ and corresponding roughly to illumination brightness, is reduced by the low-frequency attenuation thus achieving dynamic range compression while the rapidly varying component b', equal to $a(\alpha l'^*+\beta m'^*+\gamma s'^*)$ and corresponding roughly to reflectance brightness, is enhanced by the high frequency emphasis. This processing is the equivalent for color images to Stockham's work on black and white images. Notice that according to our model we should not be changing the chromatic content of the original image. Results are shown in figure 4.9. The radial frequency response used is shown in figure 4.4.

## 4.4 Chromatic processing

If we now process chromatic information and leave brightness information untouched, supposing again that the processing is linear, we can rewrite equations 4.1a to 4.1c as

$$\hat{L}^* = (\hat{L}^{1*}+\hat{L}^{2*})/2+(A-(\hat{A}^1+\hat{A}^2)/2)/a+(\beta/2)(\hat{C}_1^1-\hat{C}_1^2)+$$
$$(\gamma/2)(\hat{C}_2^2-\hat{C}_2^1) \qquad (4.4a)$$

$$\hat{M}^* = (\hat{M}^{1*}+\hat{M}^{2*})/2+(A-(\hat{A}^1+\hat{A}^2)/2)/a+((\beta-1)/2)(\hat{C}_1^1-\hat{C}_1^2)+$$
$$(\gamma/2)(\hat{C}_2^2-\hat{C}_2^1) \qquad (4.4b)$$

$$\hat{S}^* = (\hat{S}^{1*}+\hat{S}^{2*})/2+(A-(\hat{A}^1+\hat{A}^2)/2)/a+(\beta/2)(\hat{C}_1^1-\hat{C}_1^2)+$$
$$((1-\gamma)/2)(\hat{C}_2^1-\hat{C}_2^2) \qquad (4.4c)$$

where $\wedge_1$ refers to the filtering in the $C_1$-channel and $\wedge_2$ refers to the filtering in the $C_2$-channel. The meaning of those equations is not immediatly obvious but if we suppose for simplicity that the linear filtering on $C_1$ and $C_2$ is the same, then equations (4.4a) to (4.4c) rewrite very simply as

$$\hat{L}^* = \hat{L}^{1*} + (A - \hat{A}^1)/a \qquad (4.5a)$$

$$\hat{M}^* = \hat{M}^{1*} + (A - \hat{A}^1)/a \qquad (4.5b)$$

$$\hat{S}^* = \hat{S}^{1*} + (A - \hat{A}^1)/a \qquad (4.5c)$$

where, for example, $\hat{L}^{1*}$ represents the filtered version of $L^*$ by the "chromatic" filter. So the net effect is in this case the same as performing the same linear filtering on $L^*$, $M^*$, $S^*$ as on $C_1$ and $C_2$ and have it followed by masking on the resulting tristimulus values.

A complementary interpretation of this processing is the "homomorphic" one. Again we model L, M, S as products of slowly and rapidly varying components which makes the perceptual vector $\vec{P}$ the sum of two vectors $\bar{p}$ and $p'$, the coordinates of which are respectively slowly and rapidly varying. We can interpret the slowly varying part of $C_1$ and $C_2$ as the illumination chromaticity and the rapidly varying part as the reflectance chromaticity. Then a linear space invariant filtering on $C_1$ and $C_2$ characterized by a low frequency attenuation and a high frequency emphasis will enhance

saturation of objects which tend to be small and discard what may be called the chromatic component of the illumination thus realizing an automatic color balance control. This can particularly well be seen when comparing figure 4.2 with figure 4.10b where the slightly yellowish low-frequency chromatic component has been completely removed by the filtering performed on the $C_1$ and $C_2$ signals. Examples of such a process are shown in figure 4.10. The radial frequency responses used are shown in figure 4.11.

## 4.5 Brightness and chromatic processing

Finally one might want to combine brightness and chromatic enhancement as shown in figure 4.12. If we denote by $\tilde{A}$ the processed brightness, equations 4.5a to 4.5c can be rewritten as

$$\hat{L}^* = \hat{L}'^* + (\tilde{A} - \hat{A}')/a$$

$$\hat{M}^* = \hat{M}'^* + (\tilde{A} - \hat{A}')/a$$

$$\hat{S}^* = \hat{S}'^* + (\tilde{A} - \hat{A}')/a$$

again we assume for simplicity that the linear filtering on $C_1$ and $C_2$ is the same. Note that the masking is different from the one in the case of chromatic processing only. The "homomorphic" interpretation is the same as before but now a combination of paragraphs 3 and 4.

# CHAPTER 5

## APPLICATION OF THE MODEL TO IMAGE TRANSMISSION AND CODING

### 5.1 The use of the model for defining a distortion measure between color images

Shannon's *rate-distortion theory* [19, 45] was originally developed to handle such problems as efficient encoding of images and speech. Since we are interested here in images, let us review very briefly the essence of his results in that case.

Let $I(x,y,\lambda)$ be an original image that we want to transmit over some noisy channel and $\tilde{I}(x,y,\lambda)$ the received image. Let also $d(\ ,\ )$ be a positive real valued function of two images and consider $d(I,\tilde{I})$ to be the distortion that occurs when $I$ is transmitted and $\tilde{I}$ received. The performance of the system is then measured by

$$d^* = E[d(I,\tilde{I})]$$

where the expected value $E$ is taken over the ensemble of images of interest. Shannon's rate-distortion function $R(d^*)$ is a lower bound on the transmission rate required to achieve average distortion $d^*$. Moreover, Shannon's coding theorem also states that one can design a code with rate only negligibly greater than $R(d^*)$ which achieves average distortion $d^*$. Thus this function $R(d^*)$ exactly specifies the minimum achievable transmission rate $R$ required to transmit an image with average distortion level $d^*$ and provides an absolute yardstick against which to compare the performance of any practical system (see for example [19]).

There are three reasons why this potential value has not been realized to date. The first one is that there does not currently exist any tractable mathematical models for an image source, the Gaussian one being obviously a poor choice. The second one is the difficulty of computing the rate-distortion function for other than Gaussian sources and square-error distortion measures. The third one is that a distortion measure $d(\ ,\ )$ in agreement with subjective evaluation of image quality is not known. By this we mean that if $I(x,y,\lambda)$ is a reference original image and $\hat{I}_j(x,y,\lambda)$, $j = 1,2,\ldots,$ is any set of reproduced images, then $d(I,\hat{I}_j)$ should rank the reproduced images in the same order as the end user of the images.

It has been argued (Sakrison and Mannos [43]) that this might be the prime reason of the three why rate-distortion theory is not currently applicable. Some work toward the definition of such a distortion measure for black and white images has been carried out after the pioneering work of Stockham [50, 51] by Mannos [33], Rom [39] who used the mono-frequency channel model of human vision. This work is currently being extended toward including the existence of several frequency channels [2, 44]. On the side of color images, very little has been done and part of this work is an attempt to define such a distance between color images.

We will assume for simplicity that the original image is represented in some set of primaries by its corresponding tristimulus values $R(x,y)$, $G(x,y)$ and $B(x,y)$. Our reference original will thus be

$$\bar{I}(x,y) = [R(x,y),G(x,y),B(x,y)]^t$$

while the distorted version will be

$$\vec{f}(x,y) = [\tilde{R}(x,y),\tilde{G}(x,y),\tilde{B}(x,y)]^t$$

We know that

$$d(\vec{f},\vec{\hat{f}}) = \int\int ([\vec{f}(x,y)-\vec{\hat{f}}(x,y)].[\vec{f}(x,y)-\vec{\hat{f}}(x,y)])^{1/2}\,dxdy \qquad (5.1)$$

,where . stands for inner-product, is in very poor agreement with subjective evaluation. But we also know a great deal about about how the visual system processes the information it receives from the outside world and thus what is important for the human observer. In other words it seems very likely that

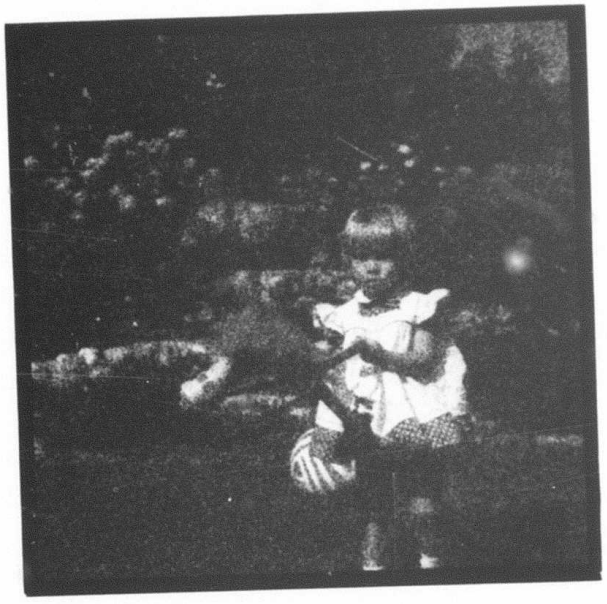
$$d(\vec{f},\vec{\hat{f}}) = \int\int ([\vec{p}^*(x,y)-\vec{\hat{p}}^*(x,y)].[\vec{p}^*(x,y)-\vec{\hat{p}}^*(x,y)])^{1/2}\,dxdy \qquad (5.2)$$

is a much better measure of the distortion between $\vec{f}$ and $\vec{\hat{f}}$ than equation (5.1). In equation (5.2) $\vec{p}^*$ and $\vec{\hat{p}}^*$ are the perceptual vectors corresponding respectively to $\vec{f}$ and $\vec{\hat{f}}$. This equation is nice because it measures distances in the $(A^*,C_1^*,C_2^*)$ space which is the $(A,C_1,C_2)$ space after spatial filtering by the linear filters of frequency responses H, $H_1$, $H_2$ (see chapter 3). This means that we are taking into account both the fact that the $(A,C_1,C_2)$ space is uniform in terms of perception (see chapter 2) and that different sensitivities to spatial frequencies exist in the three channels.

## 5.2 Agreement between distance and perceptual ranking

We artificially distorted an original image $\vec{f}$ in ways to be described. This resulted in a set of five disturbed versions $\vec{\hat{f}}_j$, $j = 1,2,\ldots,5$, of $\vec{f}$. For each $\vec{\hat{f}}_j$ we computed $d(\vec{f},\vec{\hat{f}}_j) = d_j$ according to equation (5.2). We found that the way most observers ranked the images in terms of similarity to the original was the same as the one obtained by ordering the numbers $d_j$ (the smallest number corresponding to the picture judged closest to the original).

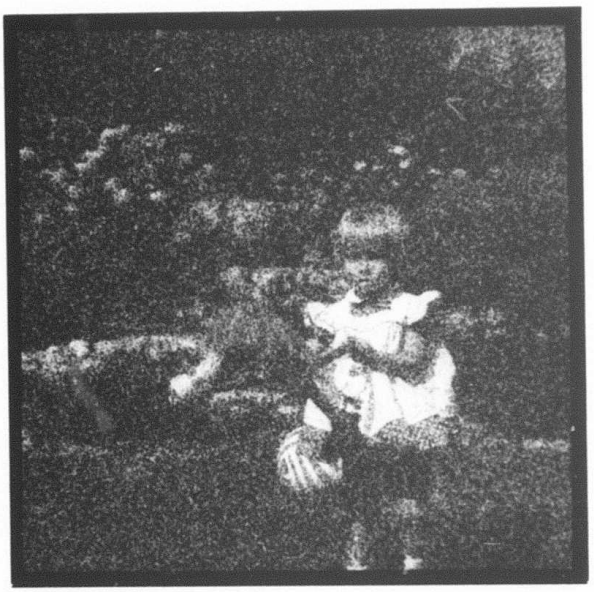
(a)



(b)

Fig. 5.1- Distorted versions of (a) "BECKY" and (b) "CAR-PORT". Noise has been added to the components corresponding to the red, green, blue primaries.

(a)



(b)

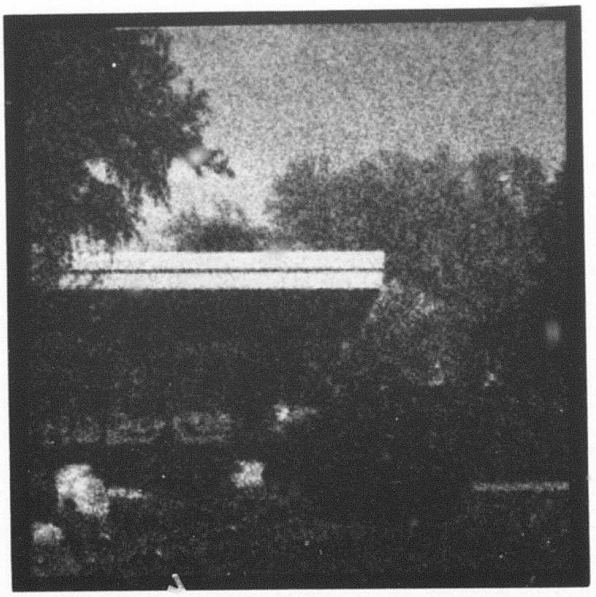
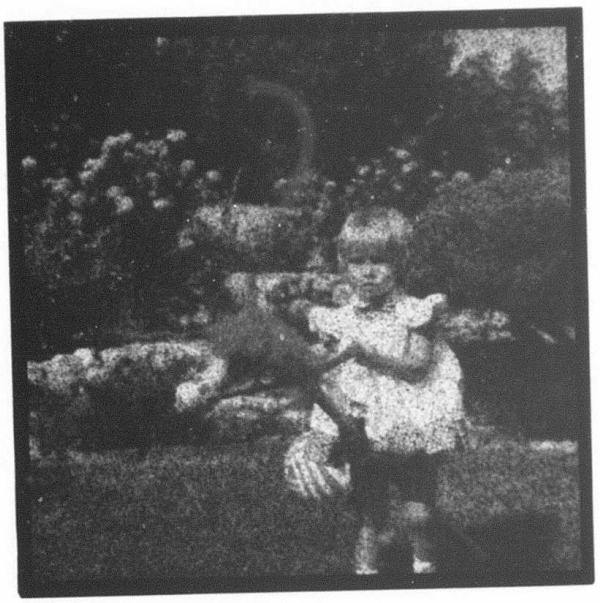


Fig. 5.2- Distorted versions of (a) "BECKY" and (b) "CAR-PORT". Noise has been added to the components corresponding to the Y, I, Q N.T.S.C. transmission primaries.

(a)



(b)

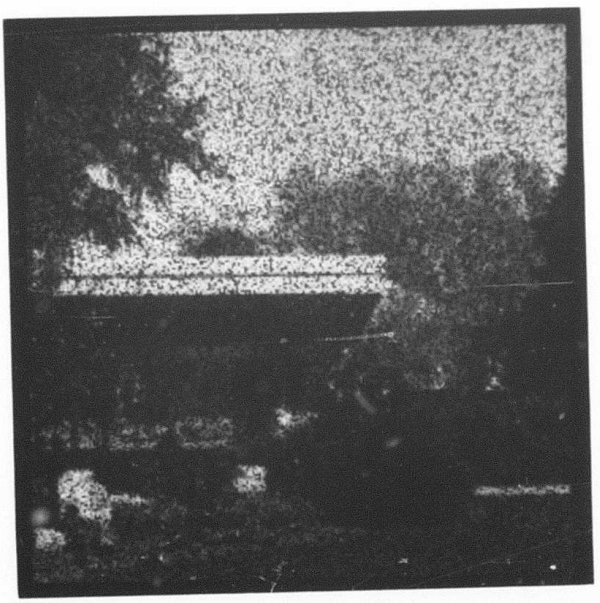
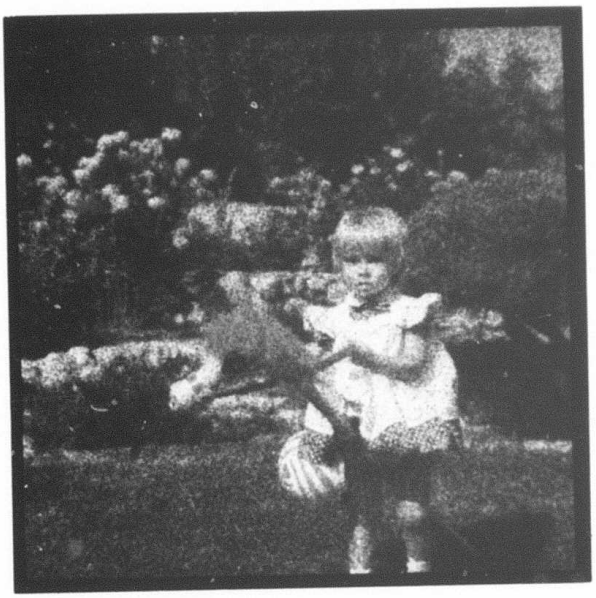


Fig. 5.3- Distorted versions of (a) "BECKY" and (b) "CAR-PORT". Noise has been added to the L*, M*, S* cones log-outputs in the model of figure 4.3.

(a)



(b)

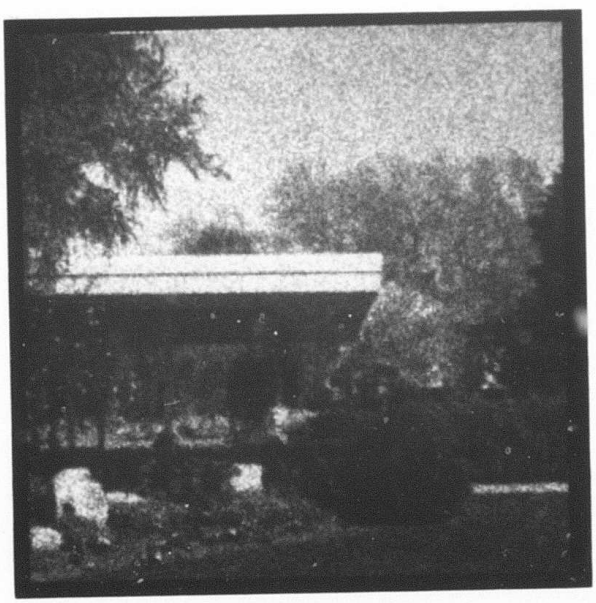


Fig. 5.4- Distorted versions of (a) "BECKY" and (b) "CAR-PORT". Noise has been added to the A, $C_1$, $C_2$ components in the model of figure 4.3.

(a)

(b)

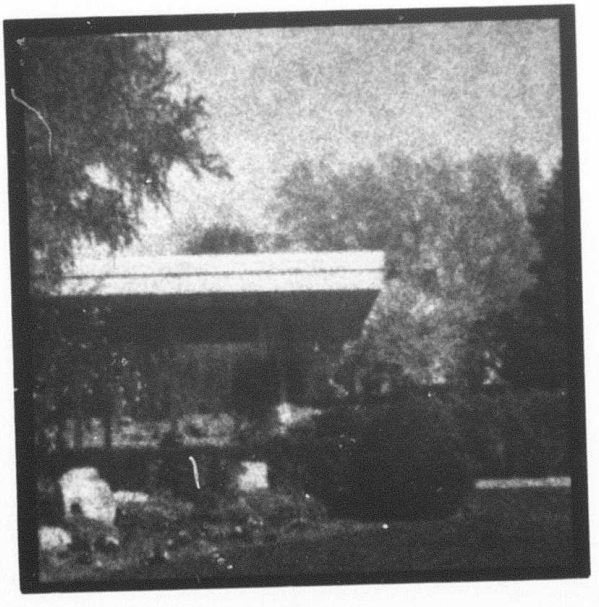


Fig. 5.5- Distorted versions of (a) "BECKY" and (b) "CAR-PORT". Noise has been added to the A', $C_1$', $C_2$' components, filtered versions of the A, $C_1$, $C_2$ components in the model of figure 4.3.

The distorted versions of "CAR-PORT" and "BECKY" numbered from 1 to 5 are shown in figures 5.1 to 5.5. The computed distortions $d_j$, (j=1,...,5), normalized to an average distortion per point, are shown in table 5.1 for "BECKY", in table 5.2 for "CAR-PORT". They correspond to the entry $D_1$. We can see that in both cases the ordering is 5, 4, 3, 1, 2. For "CAR-PORT", for example, the most disturbed image $\vec{f}_5$ is at an average distance of 20 perceptual units from the original (remember that a distance of 1 corresponds to a just noticeable difference in the $(A,C_1,C_2)$ space). $\vec{f}_1$ is 5% closer, $\vec{f}_4$ 15% closer, $\vec{f}_3$ 30% closer and $\vec{f}_2$ 35% closer.

This can be considered as a good confirmation of our claim that the distance defined by the model is "good" in the sense of paragraph 5.1. It is interesting here to make a remark on the nature of the distance or metric we defined on the $(A,C_1,C_2)$ space. We chose it to be a square-error or euclidian distance because, as we mentioned in paragraph 5.1, it is the one which mathematically is the more tractable. But we must stress the fact that from the physiologcal standpoint it is a little bit difficult to defend because it assumes some kind of mechanism able to sense the outputs of the three channels (which is not unrealistic) and somehow to compute the square-root of the sum of their squares (which is unrealistic). Another distance which would be more in agreement with what we know about neural networks can be informally defined as "the channel that shouts the loudest determines the response" and formally as the "maximum" distance whereby if O is the origin of the perceptual space and Q the point of coordinates $(A,C_1,C_2)$ then

$$OQ = \mathrm{Max}(|A|, |C_1|, |C_2|)$$

where | | stands for magnitude of. Another distance, also more defendable from the physiological standpoint, and which alike the euclidian one yields the interpretation of OQ as a measure of the total activity in all three channels would be

$$OQ = |A| + |C_1| + |C_2|$$

For those two distances, the unit spheres locus of the points representing colors which are just noticeably different from a given color would be unit cubes and the Mac-Adam ellipses would map onto unit-squares. It is pleasant to notice, as shown in tables 5.1 and 5.2, that if those distances are used for computing the distortions $d_j$ instead of the euclidian one the resulting ranking is the same (the "maximum" distance corresponds to the entry $D_7$, the "sum of the magnitudes" distance corresponds to the entry $D_8$).

Let us mention for completeness the fact that these three norms are special cases of a broad class of vector norms, the Holder-norms where

$$OQ = (|A|^\alpha + |C_1|^\alpha + |C_2|^\alpha)^{1/\alpha}$$

The sum of the magnitude norm corresponding to $\alpha = 1$, the euclidian norm corresponding to $\alpha = 2$ and the maximum norm corresponding to $\alpha = \infty$.

Obviously more work needs to be done in order to test further the metric of the $(A, C_1, C_2)$ space but those results are extremely encouraging.

## 5.3 Color image transmission and coding

The description of the way images represented in figures 5.1 to 5.5 have been distorted will provide us with a transition to

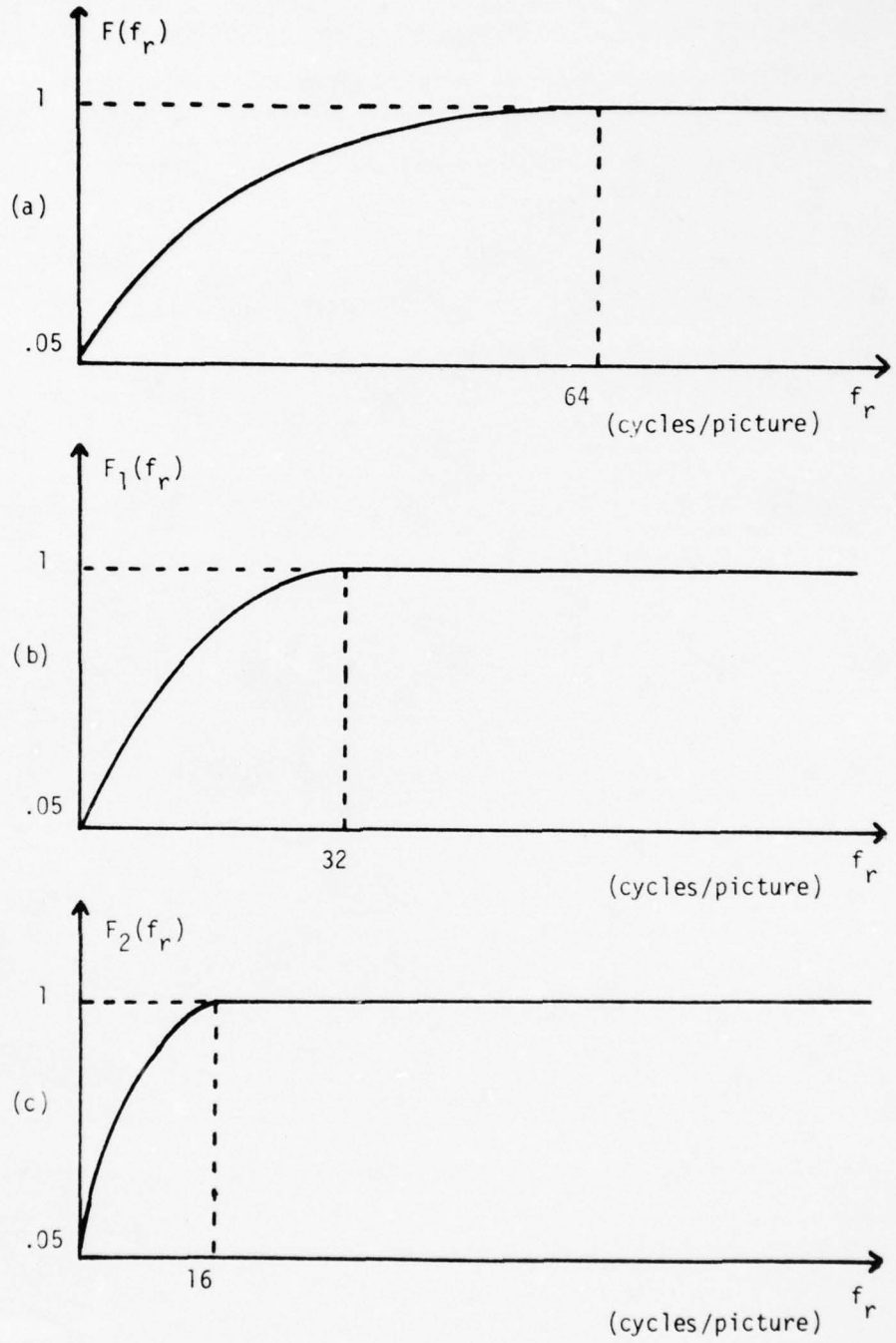



Fig. 5.6- Cross-section of the frequency responses of the filters used to obtain (a) A' from A, (b) $C_1$' from $C_1$, (c) $C_2$' from $C_2$. In each case the low-frequency part is the same as for the curves shown in figures 3.9a, 3.9c, 3.9e, respectively.

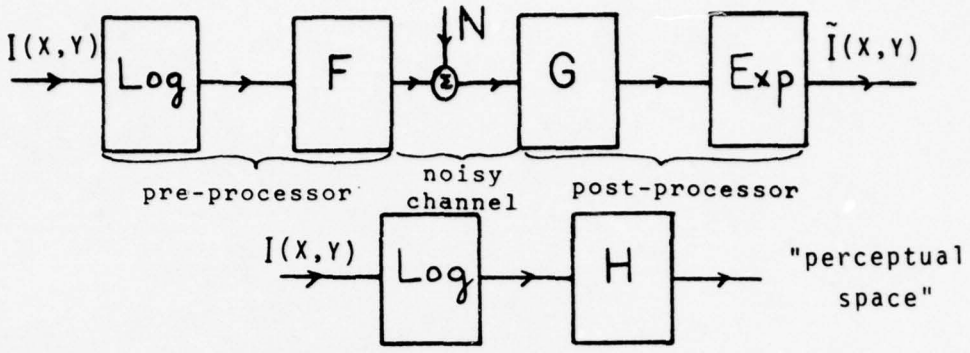



Fig. 5.7- Rom's result for transmission of an achromatic image over a noisy channel when the error criterion is a square error in the "perceptual" space.

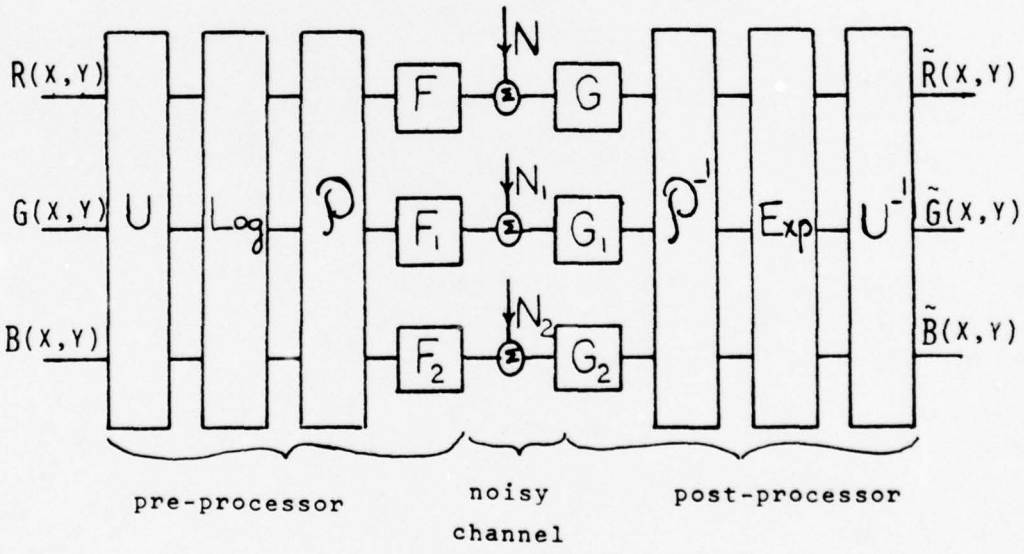



Fig. 5.8- Generalization of Rom's result to color image transmission.

the problem of image coding. Noise was added to the three components of the images in different domains. The domains were the red, green, blue display primaries for figure 5.1, the Y, I, Q NTSC transmission primaries (not gamma corrected) for figure 5.2, the L*, M*, S* cone Log-outputs for figure 5.3, the A, $C_1$, $C_2$ components of figure 3.2 for figure 5.4 and the A', $C_1$', $C_2$' components for figure 5.5. A', $C_1$', $C_2$' were obtained by filtering A, $C_1$, $C_2$ with linear filters of frequency responses equal to H, $H_1$ and $H_2$ of chapter 3 in the low frequency part up until the maximum and then flat at higher frequencies (high-pass filters) as shown in figure 5.6. The reasons for this choice will be explained later. After the noise was added the image was mapped back, if necessary, into the display primaries system and displayed.

In all cases the noise added to the three components was uncorrelated from one component to the other and for each component it was zero-average uniformly distributed white noise. Furthermore the peak to peak signal to noise ratio was 4:1 in the case of CAR-PORT and 2:1 in the case of BECKY.

If we recall that uniformly quantizing a signal with n bits is approximately equivalent to adding uniformly distributed white noise uncorrelated to the signal with peak to peak signal to noise ratio $2^n$, it is obvious by looking at figures 5.1 to 5.5 that with the same number of bits or equivalently the same relative noise level the quality achieved is much better if we use the perceptual signals A, $C_1$, $C_2$ or A', $C_1$', $C_2$'.

More formally, it has been shown recently [39] that in the case of transmission of black and white images over a noisy channel, if

one wants to minimize the mean square error in the perceptual space, the optimum pre- and post-processors are as shown in figure 5.7. F and G are the frequency responses of two spatially invariant linear filters and H is the brightness-channel frequency response of chapter 3. F depends on H, I(x,y) and the noise N, G is the corresponding Wiener-filter. As pointed out by Rom, the dependency on H is weak and in all practical cases F is a high-pass filter and G a low-pass filter.

This can be immediatly generalized to the case of color images as shown in figure 5.8. So, according to this theoretical argument, the example where we disturbed A', $C_1$' and $C_2$' corresponds almost to the optimum since the filters of figure 5.6 are high-pass filters and thus their inverses are low-pass filters. The results should indeed look better to a human observer than results obtained by different methods.

These arguments strongly support the validity of our model to define a distortion measure between color images and the more subtle claim that when processing images we should process them after a mapping so that the processed quantities are as close as possible to the perceptually important quantities to a human observer.

As a final example we would like to show applications of this model to image coding. The idea was to code the A, $C_1$, $C_2$ perceptual components of an image independently and use the fact that they are spatially filtered in different ways by the visual system. We noticed in chapter 3 that the visual acuity for the red-green channel seemed to be about twice poorer than the visual acuity for brightness and twice better than the visual acuity for

TABLE 5.1

PERCEPTUAL DISTANCES FOR "BECKY"

| | Image number | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| $D_1$ | 19.61 | 23.04 | 16.18 | 13.21 | 12.59 |
| $D_2$ | 17.60 | 21.15 | 14.18 | 11.04 | 10.35 |
| $D_3$ | 26.77 | 30.59 | 22.90 | 19.53 | 18.87 |

TABLE 5.2

PERCEPTUAL DISTANCES FOR "CAR-PORT"

| | Image number | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| $D_1$ | 18.99 | 19.67 | 16.28 | 13.15 | 12.57 |
| $D_2$ | 16.79 | 17.56 | 14.10 | 10.77 | 10.08 |
| $D_3$ | 26.41 | 27.00 | 23.47 | 19.85 | 19.22 |

(a)

(b)

Fig. 5.9- Coded versions of (a) "BECKY" and (b) "CAR-PORT" corresponding to an average bit rate of 1 bit/pixel.

the yellow-blue channel. These properties were taken into account in the encoding process where we used a spatial transform encoding method developed by Tescher [53]. Magnitude and phase of the Fourier transforms of filtered versions of A, $C_1$, $C_2$ were encoded in such a way that the average bit-rate was 0.7 bit/pixel for the brightness information, 0.2 bit/pixel for the red-green information and 0.1 bit/pixel for the yellow-blue information thus yielding a total average bit rate of 1 bit/pixel. The image was then reconstructed through the inverse of the model and displayed. The results are shown in figure 5.9. Since the originals were scanned in with 27 bits/pixel, that is 9 bits/pixel for the red, green and blue information, the reduction factor is 27:1.

# CHAPTER 6

# CONCLUSIONS

## 6.1 Psychophysics and modelling

Using the homomorphic modelling approach, we have been able to show experimentally that spatial linear inhibition is present in the chromaticity channels of the human visual system. Of course, our results hold only if the model we hypothesized is correct since the design of the entire experiment described in chapter 3 is based upon the inversion of that model.

In particular, one may argue that what we called brightness does not correspond to perceptual brightness and that, consequently, the simultaneous contrast effects we measured were partly or solely due to brightness contrast. But this criticism supposes one thing, that is that there exists a good quantitative definition of brightness. We know this is true for monochromatic lights. It is defined by the CIE $V(\lambda)$ function with the necessary corrections in the blue end of the spectrum . We were not operating with that kind of stimuli since the lights emitted by the phosphors are far from being monochromatic. An alternate possibility would have been to keep the luminance constant where the luminance Y of a stimulus of radiance $R(\lambda)$ is defined as:

$$Y = \int R(\lambda)V(\lambda)\, d\lambda$$

But this in turn assumes that Abney's law is valid which we know to

be wrong [23], especially for direct brightness matches which are the experimental conditions we were operating under.

Thus, our definition of brightness is not worse than any other definition. Furthermore, it is based on solid neurophysiological evidence and we were able to show that it predicted a relative luminous efficiency function $V^*(\lambda)$ in extremely close agreement with other experimental data. To put it another way, our definition of brightness is right in the only case where there is agreement among the scientific community on its definition, namely in the case of monochromatic lights. As a summary, we will say that the results of our psychophysical experiment are in full agreement with our model: the constant (model) brightness patterns we used were indeed judged as so by the subjects. Those results are also compatible with our hypothesis of linear lateral inhibition and summation in the chromaticity channels which, we would like to stress it again, was fundamental in the design of the experiment.

We will not claim to have solved all the problems of color vision but we hope to have contributed to a better understanding of the questions of chromatic adaptation and simultaneous color contrast effects. At any rate, this field is a most difficult one to work in because, to quote Rushton, "whereas nearly all the phenomena of nature are simply observed, those of sensory physiology can only be experienced. So in colour vision we perceive the essential hollowness of formal scientific explanation." And "the content of our hollow scientific structure is such stuff as dreams are made on: there is nothing either green or grey but thinking makes it so."

## 6.2 Image processing

After the model was calibrated using our experimental results as well as those of others, it suggested several interesting ideas. First we showed how the perceptual space offered a meaningful way of thinking about important perceptual parameters such as brightness, hue, saturation and a concise formalism to describe them quantitatively. Second we showed that the idea of structuring the perceptual space as a vector space where vector addition corresponds to tristimulus values multiplication, coupled with the existence of spatial filtering in that space accounted for contrast and constancy effects and suggested interesting image processing methods which we showed to be indeed succesful both for enhancement and transmission and coding. Third we showed that the introduction of a norm on that space, which allowed us to measure distances herein, was the key to defining a distortion measure in agreement with perceptual evaluation.

## 6.3 Further research

From the psychophysiological standpoint, one may try to verify the linearity of neural interactions for more complicated patterns than the one which were used in this study. True two-dimensional patterns could be used on $C_1$ and $C_2$ (we used only one-dimensional one), patterns exciting both $C_1$ and $C_2$ and patterns exciting all three channels could also be used.

It is also likely that more than three quantities govern our perception of colors [15] in complex scenes. We defined a new one, which we called strength, that was related to the total activity in

all three channels but did not investigate its role in perception any further. A more systematic study of such parameters may open new avenues in the fields of psychophysics and image processing.

But we think that the real improvement of this model and similar ones will come from a study of visual perception near edges. Indeed, Baudelaire [1] showed that the predictions of the homomorphic model for achromatic vision broke down for patterns with sharp edges. Of course this is also true of this model which is based on the same ideas of modelling vision as a cascade of a nonlinear memoryless stage (Log) and a linear system. The trouble is that edges determine our perception of contours and thus of shapes. That is why we think that if the next step toward a better understanding of the information processing capability of the human visual system is a step in the direction of trying to describe how we perceive and recognize shapes and objects, the problem of edge perception and the fundamental nonlinearity it introduces in the vision process will have to be solved.

We are still far from being able to explain how we structure and understand what we see with what we know about the visual system but it is hoped that studies like ours are in the right direction.

## APPENDIX A

## DERIVATION OF THE RELATIVE LUMINOUS EFFICIENCY FUNCTION

To determine experimentally the relative luminous efficiency function one can use the so-called step by step method where two patches of monochromatic lights are viewed side by side. The wavelengths of the two patches are slightly different and the radiance of one is varied until the total perceptual difference between the two patches reaches a minimum. This procedure is then repeated step by step along the spectrum.

Let us now see how such a function can be derived from the model. Let P an P' be two patches of light of tristimulus values (L,M,S) and (L',M',S') and chromaticity coordinates (l,m,s) and (l',m',s') where

$$l = L/(L+M+S) = L/\Sigma \qquad l' = L'/(L'+M'+S') = L'/\Sigma'$$

and four other similar expressions for m, m', s, s'.

We are going to make $\Sigma'$ vary while keeping l', m', s' fixed (change the radiance of patch P' without changing its color). Patches P and P' are represented in the $(A,C_1,C_2)$ space by two points Q and Q' of coordinates

$$A = a(\alpha Log(\Sigma l) + \beta Log(\Sigma m) + \gamma Log(\Sigma s))$$

$$C_1 = u_1 Log(l/m)$$

$$C_2 = u_2 Log(l/s)$$

for Q and

$$A' = a(\alpha Log(\Sigma' l') + \beta Log(\Sigma' m') + \gamma Log(\Sigma' s'))$$

$$C_1' = u_1 Log(l'/m')$$

$$C_2' = u_2 Log(l'/s')$$

for Q'. The perceptual difference D between those two patches is defined as the norm of the vector QQ'

$$D^2 = (A - A')^2 + (C_1 - C_1')^2 + (C_2 - C_2')^2$$

We want to minimize $D^2$ when varying $\Sigma'$, all other variables being constant

$$d(D^2)/d\Sigma' = d((A - A')^2)/d\Sigma 7 = -2(A - A')(dA'/d\Sigma')$$

It is interesting at this point to note that minimizing $D^2$ is equivalent to minimizing $(A - A')^2$ which is the brightness difference between the two patches. A straightforward computation yields

$$d(D^2)/d\Sigma' = (2a^2/\Sigma')(\alpha+\beta+\gamma)(\alpha Log(\Sigma' l'/\Sigma l)+\beta Log(\Sigma' m'/\Sigma m)+\gamma Log(\Sigma' s'/\Sigma s))$$

Keeping only first order terms yields

$$d(D^2)/d\Sigma' = (2a^2/\Sigma)(\alpha + \beta + \gamma)(\alpha\Delta L/L + \beta\Delta M/M + \gamma\Delta S/S)$$

Where $\Delta L = (\Sigma' l' - \Sigma l)$ and two other similar expressions for $\Delta M$ and $\Delta S$. The minimum perceptual difference is obtained for

$$\alpha\Delta L/L + \beta\Delta M/M + \gamma\Delta S/S = 0 \qquad (A.1)$$

Let us now take into account the fact that patches P and P' are monochromatic of wavelengths $\lambda$ and $\lambda + d\lambda$ and radiances $R(\lambda)$ and $R(\lambda) + dR(\lambda)$. $l(\lambda)$, $m(\lambda)$ and $s(\lambda)$ being the cone absorption curves we have

$$L = l(\lambda)R(\lambda) \qquad \Delta L = ((dl(\lambda)/d\lambda)R(\lambda) + l(\lambda)(dR(\lambda)/d\lambda))\Delta\lambda$$

and four other similar expressions.

Then

$$\Delta L/L = ((1/l(\lambda))(dl(\lambda)/d\lambda) + (1/R(\lambda))(dR(\lambda)/d\lambda))\Delta\lambda$$

and two similar expressions.

Equation (A.1) can now be rewritten as

$$(\alpha + \beta + \gamma)(1/R(\lambda))dR(\lambda)/d\lambda + (\alpha/I(\lambda))dI(\lambda)/d\lambda +$$

$$(\beta/m(\lambda))dm(\lambda)/d\lambda + (\gamma/s(\lambda))ds(\lambda)/d\lambda = 0$$

or

$$-(dR(\lambda)/R(\lambda)) = (1/\alpha + \beta + \gamma)(\alpha dI(\lambda)/I(\lambda) + \beta dm(\lambda)/m(\lambda) +$$

$$\gamma ds(\lambda)/s(\lambda))$$

This differential equation is readily integrated and yields

$$1/R(\lambda) = I(\lambda)^{\alpha/\alpha+\beta+\gamma} m(\lambda)^{\beta/\alpha+\beta+\gamma} s(\lambda)^{\gamma/\alpha+\beta+\gamma}$$

and thus the relative luminous efficiency function predicted by the model is

$$V^*(\lambda) = k/R(\lambda) = kI(\lambda)^{\alpha/\alpha+\beta+\gamma} m(\lambda)^{\beta/\alpha+\beta+\gamma} s(\lambda)^{\gamma/\alpha+\beta+\gamma}$$

where the constant k is adjusted to make the maximum equal to 1. Constants $\alpha$, $\beta$, $\gamma$ are adjusted so that this maximum occurs at 555nm.

## APPENDIX B

## OPTIMIZATION OF a, $u_1$, $u_2$

Color matches have a normal distribution around a given color center in the CIE $(x,y,0.2\log_{10}(Y))$ space (Brown and Mac-Adam 1949 [4]). Given that just-noticeable differences are about three times standard deviations, the locus of colors just noticeably different from the given color center is an ellipsoid. We would like this ellipsoid to map to a sphere of unit radius in the $(A,C_1,C_2)$ space.

In order to do that, we first assume that chromaticity and brightness errors are independent which allows us to optimize independently a and $(u_1,u_2)$.

For an achromatic color, we have L=M=S and thus A=aLog(L). Also, if the model is calibrated for the equal energy white, then we also have L=X=Y=Z and thus A=aLog(Y). This implies

$$(dA)^2 = a^2(dY/Y)^2$$

If $(dA)^2=1$ is the just noticeable difference that we want

$$a^2(dY/Y)^2 = (60.10^4/9)(dY/11.53Y)^2$$

and thus

$$a = 22.6$$

Optimization of parameters $u_1$ and $u_2$ is done in the following way. We take the Mac-Adam ellipses, projections of the standard-deviations ellipsoids on the (x,y) plane, and map them onto the $(C_1,C_2)$ plane of the model (see figure B.1). We then adjust $u_1$

and $u_2$ to get the best fit with circles of radius 1/3. Table 1 shows the results for $u_1$=64 and $u_2$=10.

TABLE B.1

| Color center | | Mac-Adam ellipses | | | Model | |
|---|---|---|---|---|---|---|
| $x_o$ | $y_o$ | 1000u | 1000v | θ | 3U | 3V |
| 0.160 | 0.057 | 0.85 | 0.35 | 62.5 | 1.44 | 0.72 |
| 0.187 | 0.118 | 2.2 | 0.55 | 77.0 | 0.72 | 0.73 |
| 0.253 | 0.125 | 2.5 | 0.50 | 55.5 | 0.70 | 0.70 |
| 0.150 | 0.680 | 9.6 | 2.3 | 105 | 1.70 | 0.82 |
| 0.131 | 0.521 | 4.7 | 2.0 | 112.5 | 0.64 | 0.74 |
| 0.212 | 0.550 | 5.8 | 2.3 | 100.0 | 0.97 | 0.83 |
| 0.258 | 0.450 | 5.0 | 2.0 | 92.0 | 0.88 | 0.81 |
| 0.152 | 0.365 | 3.8 | 1.9 | 110.0 | 0.63 | 0.93 |
| 0.280 | 0.385 | 4.0 | 1.5 | 75.5 | 0.73 | 0.69 |
| 0.380 | 0.498 | 4.4 | 1.2 | 70.0 | 1.72 | 0.48 |
| 0.160 | 0.200 | 2.1 | 0.95 | 104.0 | 0.44 | 0.82 |
| 0.228 | 0.250 | 3.1 | 0.90 | 72.0 | 0.62 | 0.61 |
| 0.305 | 0.323 | 2.3 | 0.90 | 58.0 | 0.51 | 0.49 |
| 0.385 | 0.393 | 3.8 | 1.6 | 65.5 | 0.95 | 0.77 |
| 0.472 | 0.399 | 3.2 | 1.4 | 51.0 | 1.29 | 0.73 |
| 0.527 | 0.350 | 2.6 | 1.3 | 20.0 | 1.25 | 0.84 |
| 0.475 | 0.300 | 2.9 | 1.1 | 28.5 | 1.00 | 0.84 |
| 0.510 | 0.236 | 2.4 | 1.2 | 29.5 | 0.69 | 1.60 |
| 0.596 | 0.283 | 2.6 | 1.3 | 13.0 | 1.53 | 1.50 |
| 0.344 | 0.284 | 2.3 | 0.90 | 60.0 | 0.45 | 0.60 |
| 0.390 | 0.237 | 2.5 | 1.0 | 47.0 | 0.52 | 0.93 |
| 0.441 | 0.198 | 2.8 | 0.95 | 34.5 | 0.64 | 1.39 |
| 0.278 | 0.223 | 2.4 | 0.55 | 57.5 | 0.56 | 0.43 |
| 0.300 | 0.163 | 2.9 | 0.60 | 54.0 | 0.62 | 0.72 |
| 0.365 | 0.153 | 3.6 | 0.95 | 40.0 | 0.75 | 1.60 |

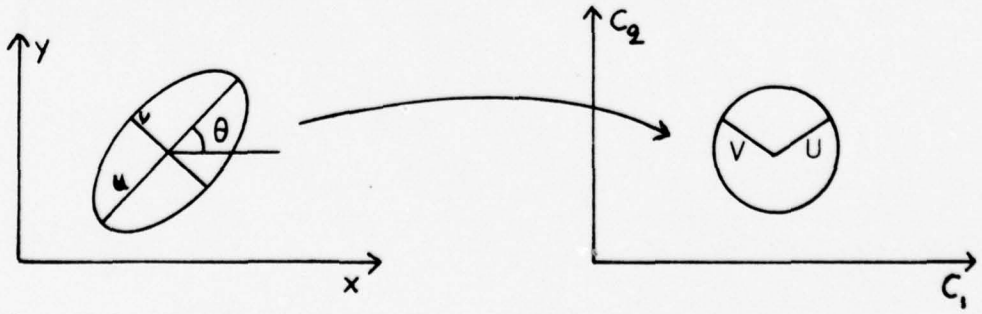



Fig. B.1- Mapping of MacAdam ellipses into circles.

APPENDIX C

IMPORTANT MATRICES

We had to solve several times the problem of finding the matrix representing a change of primaries, the standard primaries being defined by the CIE X, Y, Z tristimulus values. This situation is depicted on figures C.1 and C.2

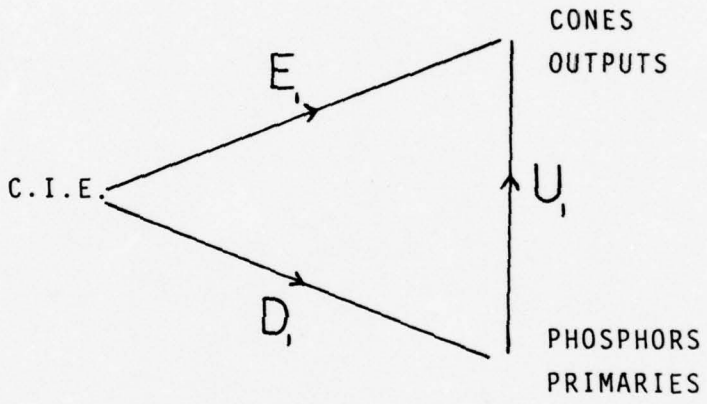



Fig. C.1-Computation of the $U_1$ matrix from matrices $E_1$ and $D_1$. where the reference white $\vec{w}_1 = [1, 1.051, 1.144]^t$ is the D6500 white.

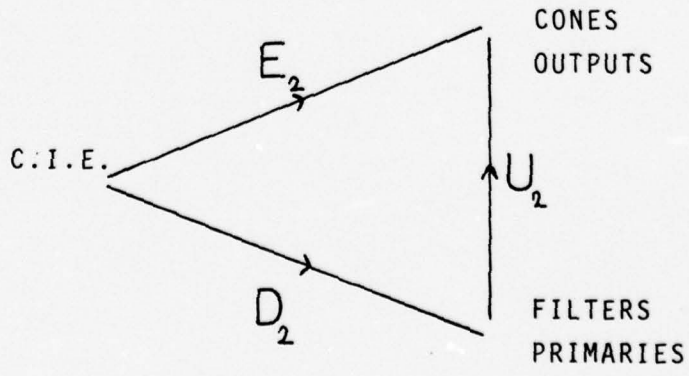



Fig. C.2- Computation of the $U_2$ matrix from matrices $E_2$ and $D_2$.

where the reference white $\vec{W}_2 = [1, 1.019, 1.174]^t$ is the CIE illuminant C.

Thus we have $U_i = E_i D_i^{-1}$ $(i=1.2)$. $D_i$ and $E_i$ are normalized so that:

$$\vec{I} = E_i \vec{W}_i \qquad (i=1,2) \tag{C.1a}$$

and

$$\vec{I} = D_i \vec{W}_i \qquad (i=1,2) \tag{C.1b}$$

$E_i$, $(i=1,2)$, stem from $E^{-1}$ given by Stiles ([61] p.515)

$$E^{-1} = \begin{bmatrix} 661 & 1260 & -112 \\ -438 & 1620 & 123 \\ .708 & 0 & 417 \end{bmatrix} \tag{C.2}$$

$D_i$ stems from $F_i$ such that $F_i^{-1}$ has for column vectors the chromaticity coordinates x, y, z of the phosphors lights

$$F_1^{-1} = \begin{bmatrix} .628 & .343 & .155 \\ .345 & .585 & .066 \\ .027 & .072 & .779 \end{bmatrix} \tag{C.3}$$

in the same way $D_2$ stems from $F_2$ such that $F_2^{-1}$ has for column vectors the chromaticity coordinates x, y, z of the light coming through the Wratten filters 25, 58 and 47B when illuminated by the CIE illuminant C

$$F_2^{-1} = \begin{bmatrix} .6808 & .2425 & .1579 \\ .3190 & .6923 & .0187 \\ .0002 & .0652 & .8234 \end{bmatrix} \tag{C.4}$$

The normalization procedure goes as follows for $E_i$ as an example. Let us rewrite $E^{-1}$ as

$$E^{-1} = [\vec{T}_1, \vec{T}_2, \vec{T}_3]$$

where $\vec{T}_i$ $(i=1,2,3)$ is the i th column vector. Then $E_i^{-1}$ is of the form

$$E_1^{-1} = [\lambda_1\vec{T}_1, \lambda_2\vec{T}_2, \lambda_3\vec{T}_3] = E^{-1}\lambda$$

where

$$\lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

and $\lambda_1$, $\lambda_2$, $\lambda_3$ are three parameters. Then equation (C.1a) rewrites as

$$\vec{W}_1 = E_1^{-1}\vec{I} = E^{-1}\lambda\ \vec{I}$$

thus

$$\lambda\vec{I} = E\ \vec{W}_1$$

or

$$[\lambda_1, \lambda_2, \lambda_3]^t = E\ \vec{W}_1$$

Matrices $D_1$, $E_2$, $D_2$ are computed exactly the same way. One finally has:

$$U_1 = \begin{bmatrix} .2457 & .6840 & .0703 \\ .1101 & .7625 & .1273 \\ .0132 & .0842 & .9026 \end{bmatrix}$$

$$U_2 = \begin{bmatrix} .3634 & .6102 & .0264 \\ .1246 & .8138 & .0616 \\ .0009 & .0602 & .9389 \end{bmatrix}$$

Another important matrix is the matrix which achieves the separation of brightness and chromatic information

$$\mathcal{P} = \begin{bmatrix} a\alpha & a\beta & a\gamma \\ u_1 & -u_1 & 0 \\ u_2 & 0 & -u_2 \end{bmatrix}$$

where a=22.6, $u_1$=64 and $u_2$=10.

# APPENDIX D

## CALIBRATION PROCEDURES

We saw in chapter 4 what the nonlinearities involved in the input process were and how they had been compensated for. The output process could be two things: either display the image on the face of a television monitor (chapter 3) or display the image row by row on the face of a CRT which was used to expose a piece of color film.

In the first case the calibration procedure consisted in two steps. First the white (corresponding to equal drive signals on the three guns) was precisely adjusted to D6500 using a split-field color comparator. Then the nonlinearity of the tube was estimated in the following way. A digital step-wedge pattern consisting of density steps between 2.5 and 8 by steps of .5 (thus corresponding to digital intensities between 2↑2.5 and 2↑8 by steps of 2↑.5) was displayed on the face of the television monitor and the luminance of every step was measured with a Tektronix J16 photometer connected to a J6503 luminance probe. The graph luminance/intensity turned out to be an exact straight line of slope 2.5 on a Log-Log scale thus indicating a gamma of 2.5 for the tube. This was compensated for by table lookup in the Comtal before conversion through the 8-bit DACS.

In the second case the method was pretty much the same except for the complication introduced by the film. A step-wedge pattern

was used to expose a piece of color film through three filters. The film was then processed and the negative printed for an exact densitometric grey on the fifth step. The D-LogE curves were then measured for the three colors red, green and blue. A typical set of such curves is shown in figure D.1. The corresponding nonlinearities (CRT and film) were again compensated for by table lookup before conversion through the 11-bit DAC. Examples of an uncompensated and properly compensated color-wedge are shown in figure D.2. For the compensated one, the three D-LogE curves are straight lines of slope 1, thus guaranteeing that photographic intensities correspond closely to digital intensities.

The procedure for printing on photographic paper the images shown in this report was then the following. Every time we exposed color film on the CRT, we also exposed a piece of film with a wedge. The films were then developed. The wedge was printed for a grey so that the three D-LogE curves were straight lines of slope 1. Then the whole batch of negatives was printed exactly the same way as the wedge thus reducing to a minimum distortions introduced by film processing in the photo-lab.

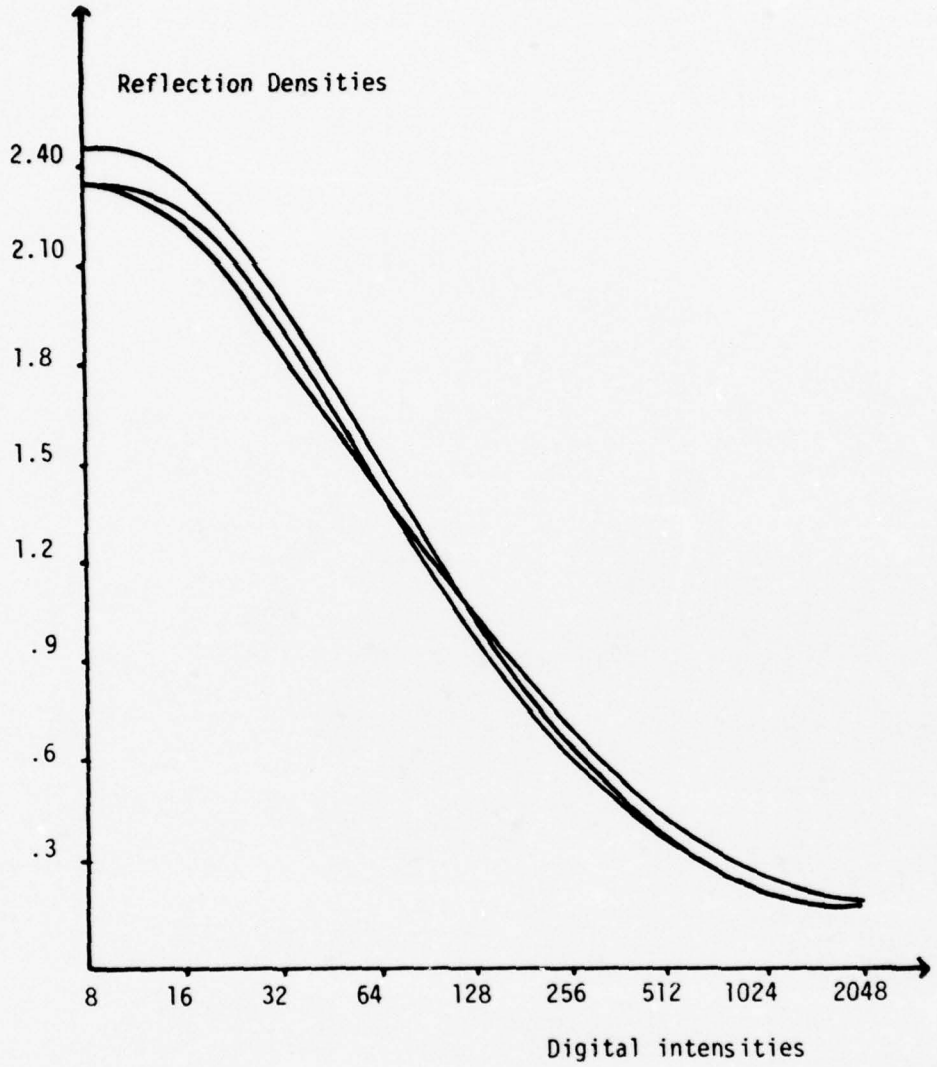
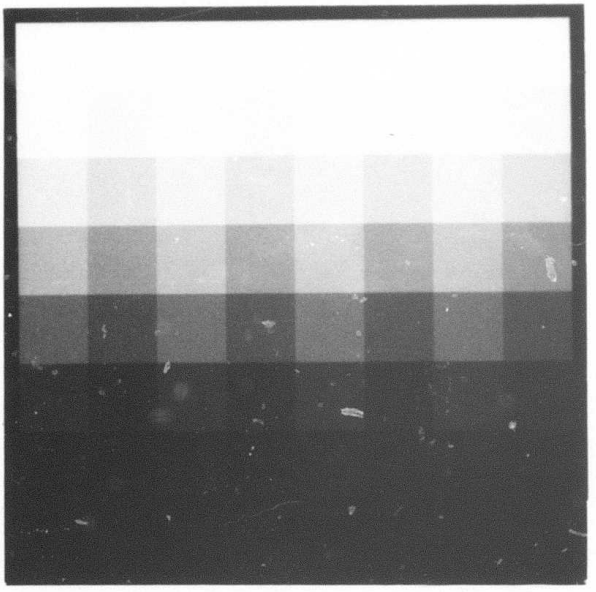



Fig. D.1- D-LogE curves measured from an uncompensated color stepwedge.

(a)



(b)

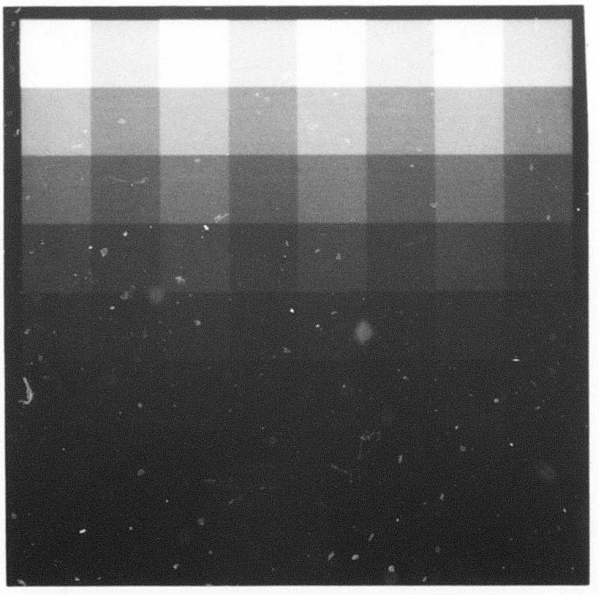


Fig. D.2- (a) uncompensated and (b) compensated color stepwedges.

# APPENDIX E

## DESCRIPTION OF THE PSYCHOPHYSICAL EXPERIMENT

### E.1 The experiment

The experiment was intended to test the linearity of spatial neural interaction in the chromaticity channels $C_1$ and $C_2$ and answer the question of whether this interaction was characterized by an attenuation of low spatial frequencies. We excited one channel at a time with a pure chromatic pattern of the form

$$C_1 = c_1 + k_1(\sin(2\pi fx) + \alpha\sin(6\pi fx)) \qquad \text{(E.1a)}$$

$$C_2 = c_2 \qquad \text{(E.1b)}$$

and

$$C_1 = c_1 \qquad \text{(E.1c)}$$

$$C_2 = c_2 + k_2(\sin(2\pi fx) + \alpha\sin(6\pi fx)) \qquad \text{(E.1d)}$$

following the method pioneered by Baudelaire. The brightness channel was constant in both cases (A=a). The values of parameters $u_1$ and $u_2$ controlled the position of the average point in the $(C_1,C_2)$ plane (see figure). After mapping through the inverse of the model of figure 2.3, that is to say through followed by exponentiation on each component and $U^{-1}$, those patterns were presented to a set of five observers of both sexes on the face of a COMTAL where they sustained a field of view of 7 degrees.

We experimented with four spatial frequencies

$f_1$ = 0.142 cycles/degree

$f_2$ = 0.284 cycles/degree

$f_4$ = 0.568 cycles/degree

$f_6$ = 1.136 cycles/degree

In all cases, all subjects declared that the brightness was uniform accross the pattern thus indicating that the brightness defined by the model is in good agreement with perceptual brightness.

Two central bands whose subjective appearance is controlled by simultaneous color contrast were indicated by markers. The subjects were asked to focus their gaze at the center of the pattern and immediately report verbally their judgement by one of the following answers:

R (right): The band to the right is redder

L (left): The band to the left is redder

S (same): The two bands are the same color

for the $C_1$-channel and

R (right): The band to the right is yellower

L (left): The band to the left is yellower

S (same): The two bands are the same color

for the $C_2$-channel.

For each spatial frequency and every channel, the parameter $\alpha$ was varied from 1 to .1 by steps of .1. The contrast was kept constant by adjusting the parameter k in equations (E.1a) and (E.1d). For $\alpha$=1 a strong simultaneous color contrast illusion was produced except for frequency $f_6$ in the $C_2$-channel. For $\alpha$=.1, the fast varying component disappears completely in equations (E.1a) and (E.1d), thus producing what may be called an over-compensation of the color illusion. The idea of the experiment is to estimate the

intermediate value of $\alpha$ which produces target bands of same color. This value of $\alpha$ is then considered as indicating the value of the relative amplification of frequencies f and 3f for the particular chromatic channel tested. Each set of ten stimuli was shown twice to every subject, the order in the set being reversed from one presentation to the next.

E.2 The results

The values $\alpha_i$ (i=1,2) for which the target bands appear to be the same color are defined as the point of subjective equality (PSE) for the i th channel. For every channel, every observer and every frequency, the parameter $\alpha_i$ was estimated by taking the average of the two presentations. For each presentation, the PSE was taken as the mean of the transition values $T^+$ and $T^-$ of $\alpha_i$ where the subjective judgement switched from R to S and from S to L respectively, as shown on the following example based on the typical data of table 1:

average of $T^+ = \langle T^+ \rangle = 0.50$

average of $T^- = \langle T^- \rangle = 0.40$

interval of uncertainty $IU = \langle T^+ \rangle - \langle T^- \rangle = 0.10$

point of subjective equality $PSE = (\langle T^+ \rangle + \langle T^- \rangle)/2 = 0.45$

Tables 2 and 3 contain the experimental results for the $C_1$-channel, tables 4 and 5 contain the experimental results for the $C_2$-channel. We averaged these results over the five subjects and obtained the following means and standard-deviations

$C_1$-channel:

point of subjective equality

| frequency | $f_1$ | $f_2$ | $f_3$ | $f_4$ |
|---|---|---|---|---|
| mean | 0.50 | 0.54 | 0.50 | 0.49 |
| s.dev. | 0.05 | 0.04 | 0.05 | 0.11 |

interval of uncertainty

| frequency | $f_1$ | $f_2$ | $f_3$ | $f_4$ |
|---|---|---|---|---|
| mean | 0.13 | 0.16 | 0.14 | 0.15 |
| s.dev. | 0.05 | 0.06 | 0.05 | 0.09 |

$C_2$-channel:

point of subjective equality

| frequency | $f_1$ | $f_2$ | $f_3$ |
|---|---|---|---|
| mean | 0.61 | 0.59 | 0.63 |
| s.dev | 0.05 | 0.12 | 0.12 |

Interval of uncertainty

| frequency | $f_1$ | $f_2$ | $f_3$ |
|---|---|---|---|
| mean | 0.10 | 0.14 | 0.10 |
| s.dev. | 0.05 | 0.04 | 0.03 |

E.3 Conclusions

We concluded from these experimental results that the point of subjective equality was about constant over the range of frequencies studied, equal to 0.50 for the $C_1$-channel and 0.60 for the $C_2$-channel. The yellow-blue channel has a lower contrast sensitivity than the red-green channel combined with a lower visual acuity.

Indeed, the results for frequency $f_4$ are not shown for this channel because the illusion was seen in reverse by four out of the five observers and seen much weaker by the fifth one. As we mentioned in chapter 3, we used this fact to estimate the position of the peak frequency of the frequency response $H_2$.

TABLE E.1

TYPICAL RESULTS FOR ONE SET

OF STIMULI FOR THE $C_1$-CHANNEL

| $\alpha$ | Presentations | |
|---|---|---|
| 1 | R | R |
| 0.9 | R | R |
| 0.8 | R | R |
| 0.7 | R | R |
| 0.6 | S | S |
| 0.5 | S | S |
| 0.4 | L | S |
| 0.3 | L | L |
| 0.2 | L | L |
| 0.1 | L | L |
| $T^+$ | .65 | .65 |
| $T^-$ | .45 | .35 |

TABLE E.2

POINT OF SUBJECTIVE EQUALITY

($C_1$-CHANNEL)

Frequencies

| | $f_1$ | $f_2$ | $f_3$ | $f_4$ |
|---|---|---|---|---|
| Subject 1 | .450 | .525 | .500 | .500 |
| Subject 2 | .575 | .575 | .575 | .600 |
| Subject 3 | .525 | .500 | .525 | .600 |
| Subject 4 | .500 | .600 | .475 | .475 |
| Subject 5 | .450 | .500 | .425 | .300 |

TABLE E.3

INTERVAL OF UNCERTAINTY

($C_1$-CHANNEL)

Frequencies

| | $f_1$ | $f_2$ | $f_3$ | $f_4$ |
|---|---|---|---|---|
| Subject 1 | .100 | .250 | .200 | .200 |
| Subject 2 | .150 | .150 | .150 | 0 |
| Subject 3 | .050 | .100 | .050 | .200 |
| Subject 4 | .150 | .100 | .150 | .250 |
| Subject 5 | .200 | .200 | .150 | .100 |

TABLE E.4

POINT OF SUBJECTIVE EQUALITY

($C_2$-CHANNEL)

Frequencies

| | $f_1$ | $f_2$ | $f_3$ |
|---|---|---|---|
| Subject 1 | .625 | .675 | .675 |
| Subject 2 | .550 | .400 | .500 |
| Subject 3 | .650 | .675 | .800 |
| Subject 4 | .675 | .700 | .700 |
| Subject 5 | .550 | .500 | .475 |

TABLE E.5

INTERVAL OF UNCERTAINTY

($C_2$-CHANNEL)

Frequencies

| | $f_1$ | $f_2$ | $f_3$ |
|---|---|---|---|
| Subject 1 | .050 | .150 | .150 |
| Subject 2 | .100 | .200 | .100 |
| Subject 3 | .100 | .150 | .100 |
| Subject 4 | .050 | .100 | .100 |
| Subject 5 | .200 | .100 | .150 |

## REFERENCES


[1] Baudelaire, P., "Digital picture processing and psychophysics: a study of brightness perception." Ph.D dissertation, University of Utah (1973).

[2] Baxter, B., "Image processing in the human visual system." Ph.D dissertation, University of Utah, December 1975.

[3] Blackwell, H.R., and O.M. Blackwell, "Rod and cone receptor mechanisms in typical and atypical congenital achromatopsia." Vision Res. 1, pp.62-107 (1961).

[4] Brown, W.R.J. and D.L. MacAdam, "Visual sensitivities to combined chromaticity and luminance differences." Journal of the Optical Society of America, vol.39, no.10, pp.808-834.

[5] Campbell, F.W., and D.G. Green, "Optical and retinal factors affecting visual resolution." Journal of Physiology 181, pp.576-593 (1965).

[6] Cornsweet, T.N., "Visual perception." New York: Academic Press, 1970

[7] Dartnall, H.J.A., "The interpretation of spectral sensitivity curves." Brit. Med. Bull. 9, 24 (1953).

[8] Davidson, M.L., "Perturbation approach to spatial brightness interaction in human vision." Journal of the Optical Society of America, vol.58, pp.1300-1308, September 1968.

[9] Davidson, M.J., and J.A. Whiteside, "Human brightness perception near sharp contours." Journal of the Optical Society of America, vol.61, pp.530-536, April 1971.

[10] De Valois, R.L., C.J. Smith, S.T. Kital, A.J. Karcly, "Responses of single cells in different layers of the primate lateral geniculate nucleus to monochromatic light." Science 127, pp.238-239 (1958).

[11] De Valois, R.L., "Analysis and coding of color vision in the primate visual system." Cold Spr. Harb. Symp. quant. Biol. 30, pp.567-579 (1965).

[12] De Valois, R.L., "Central mechanisms of color vision." Handbook of Sensory Physiology, vol.VII/3, Central Processing of Visual Information A, pp.209-253 (1973).

[13] De Valois, R.L., I. Abramov, G.H. Jacobs, "Analysis of response patterns of LGN cells." Journal of the Optical Society of America, vol.56, pp.966-977 (1966).

[14] De Valois, R.L., and P.L. Pease, "Contours and contrast: responses of monkeys lateral geniculate nucleus cells to luminance and color figures." Science vol.171, pp.694-696 (1971).

[15] Ercoles-Guzzoni, A.M., and A. Fiorentini, "Simultaneous contrast effects produced by non-uniform coloured fields." Atti Fond. Giorgio Ronchi vol.13, pp.135-144.

[16] Evans, R.M., "The perception of color." John Wiley & Sons, Inc., 1974.

[17] Frei, W., "A new model of color vision and some practical implications." USCIPI report 530, pp.128-143.

[18] Fry, G.A., "Mechanisms subserving simultaneous brightness contrast." American Journal of Optometry, vol.45, pp.1-17 (1948).

[19] Gallager, R.G., "Information theory and reliable communication." John Wiley and Sons, Inc., 1968.

[20] Graham, C.H., "Vision and visual perception." New York: John Wiley & Sons, 1965.

[21] Green, D.G., "The contrast sensitivity of the colour mechanisms of the human eye." Journal of Physiology, vol.196, pp.415-429 (1968).

[22] Green, D.G., and M.B. Fast, "On the appearance of Mach bands in gradients of varying colors." Vision Research, vol.11, pp.1147-1155, 1971.

[23] Guth, S.L., N.J. Donley, R.T. Marrocco, "On luminance additivity and related topics." Vision Research, vol.9, pp. 537-575 (1969).

[24] Hartline, H.K., "Visual receptors and retinal interaction." Science, vol.164, pp.270-278, April 1969.

[25] Hering, E., "Outlines of a theory of the light senses." Translated by Leo M. Hurvich and Dorothea Jameson. Cambridge, Mass., Harvard University Press, 1964.

[26] Hurvich, L.M., and D. Jameson, "An opponent-process theory of color vision." Psychological Review, vol.64, pp.384-404 (1957).

[27] Hurvich, L.M., and D. Jameson, "The perception of brightness and darkness." Boston: Allyn and Bacon, 1966.

[28] Jacobson, J.Z., and G.E. McKinnon, "Coloured Mach bands." Canadian Journal of Psychology, vol.23, pp.56-65, 1969.

[29] Judd, D.B., "Hue saturation and lightness of surface colors with chromatic illumination." Journal of the Optical Society of America, vol.30, pp.2-32, January 1940.

[30] Judd, D.R., "Color in business, science and industry." New York, Wiley, 1963.

[31] Koenderik, J.J., W.A. van de Grind, M.A. bouman, "Opponent color coding: a mechanistic model and a new metric for color space." Kybernetik, vol.10, pp.78-98, February 1972.

[32] Land, E.H., and J.J. McCann, "Lightness and retinex theory." Journal of the Optical Society of America, vol.61, pp.1-11, January 1971.

[33] Mannos, J.R., "A class of fidelity criteria for the encoding of visual images." Ph.D dissertation, University of California, Berkeley, December 1972.

[34] Meessen, A.,"A simple non-linear theory of color perception and contrast effect." Kybernetik, vol.4, pp. 48-54 (1967).

[35] Oppenheim, A.V., R.W. Schafer and T.G. Stockham Jr., "Non-linear filtering of multiplied and convolved signals." Proceedings of the I.E.E.E., vol.56, pp.1264-1291, August 1968.

[36] Pitt, F.H.G., "The nature of normal trichromatic and dichromatic vision." Proc. Roy. Soc. Lond. 1328, pp. 101-117 (1944).

[37] Ratliff, F., "Mach bands: quantitative studies on neural networks in the retina." San Francisco: Holden-Day 1965.

[38] Richards, W., and E.E. Parks, "Model for color conversion." Journal of the Optical Society of America, vol.61, pp.971-976, July 1971.

[39] Rom, R.J., "Image transmission and coding based on human vision." Ph.D dissertation, University of Utah, August 1975.

[40] Rushton, W.A.H, "A cone pigment in the protanope." Journal of Physiology (Lond.), vol.168, pp.345-359 (1963).

[41] Rushton, W.A.H., "A foveal pigment in the deuteranope." Journal of Physiology (Lond.), vol.176, pp.24-37 (1965).

[42] Rushton, W.A.H, "Pigments and signals in colour vision." Journal of Physiology, vol.220, pp.1-31P (1972).

[43] Sakrison, T.G., and J.L. Mannos, "The effects of a visual

fidelity criterion on the encoding of images." I.E.E.E. Transactions on Information Theory, vol.IT-20, no.4, pp.525-536, July 1974.

[44] Sakrison, T.G., M. Halter, H. Mastafavi, "Properties of the human visual system as related to the encoding of images." To appear in New Directions in Signal Processing in Communication and Control.

[45] Shannon, C.E., "A mathematical theory of communication." Bell Syst. Tech. J., vol.17, pp.623-656, October 1948.

[46] Shklover, D.A., "The equicontrast colorimetric system." Visual problems of Colour, vol.II, 605 (1957). Also in Natl. Phys. Lab. Symp., no.8, London, Her Majesty's Stationery Office, 1958.

[47] Stiles, W.L., "The directional sensitivity of the retina and the spectral sensitivities of the rods and cones." Proc. Roy. Soc. (Lond.), B127, pp.64-105 (1939).

[48] Stiles, W.L., "A modified Helmholtz line element in brightness-colour space." Proc. Phys. Soc. (Lond.), 58, pp.41-65, (1946).

[49] Stiles, W.L., "Colour vision: the approach through increment threshold sensitivity." Proc. nat. Acad. Sci. (Wash.), 45, pp.100-114 (1959).

[50] Stockham, T.G., Jr., "Image processing in the context of a visual model." Proceedings of the I.E.E.E., vol.60, Picture Processing, pp. 828-842, July 1972.

[51] Stockham, T.G., Jr., "Intra-frame encoding for monochrome images by means of a psychophysical model based on non-linear filtering of multiplied signals." Proceedings of the 1969 Symposium on Picture Bandwith Reduction, New York: Gordon and Breach, 1972.

[52] Takasaki, H., "Von Kries coefficient law applied to subjective color change induced by background color." Journal of the Optical Society of America, vol.59, pp.1370-1376, October 1969.

[53] Tescher, A.G., "The role of phase in adaptive image coding." USCIPI Report 510.

[54] Tomita, T., "Electrical response of single photo receptors." Proceedings of the I.E.E.E., vol.56, Neural Studies, pp.1015-1023, June 1968.

[55] Van der Horst, G.J.C. and M.A. Bouman, "On searching for "Mach band type" phenomena in colour vision." Vision Research, vol.7, pp.1027-1029 (1967).

[56] Van der Horst, G.J.C., "Fourier analysis and color discrimination." Journal of the Optical Society of America, vol.59, no.12, December 1969.

[57] Wallis, R.H., "Film recording of digital color images." USCIPI Report 570 (May 1975).

[58] Werblin, F.S., "The control of sensitivity in the retina." Scientific American, vol.288, no.1, pp.70-79, January 1973.

[59] Wiesel, T.N., and D.H. Hubel, "Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey." J. Neurophysiol. 29, pp.1115-1156 (1966).

[60] Willmer, E.N., "A physiological basis for human color vision in the central fovea." Doc. Ophtalm. 9, pp.235-313 (1955).

[61] Wyszecki, G., and W.S. Stiles, "Color Science." John Wiley and Sons, Inc., New York (1967).

## ACKNOWLEDGMENTS

I express my appreciation to Dr. Thomas G. Stockham, Jr. for suggesting the topic of this research and his supervision and advice until its completion. I am also grateful to Dr. Elaine Cohen for her recommendations on the content and style of my dissertation, to Dr. Steven Boll who helped in evaluating this work and to Dr. Brent Baxter for many fruitful conversations.

I want to thank also the many members of the Sensory Information Processing Group who have helped me directly or indirectly by their participation in the elaboration and maintenance of the hardware and software systems I used. Special thanks go to George Randall, Dick Warnock and Barden Smith. I am also very grateful to Mike Milochik whose expertise in the photo lab has made this research possible and to my wife Agnes for her constant support and understanding.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>UTEC-CSc-77-029 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE *(and Subtitle)*<br>Digital Color Image Processing and Psychophysics Within the Framework of a Human Visual Model | | 5. TYPE OF REPORT & PERIOD COVERED<br>Technical Report<br>6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>Olivier Dominique Faugeras | | 8. CONTRACT OR GRANT NUMBER(s)<br>DAHC-15-73-C-0363 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Computer Science Deaprtment<br>University of Utah<br>Salt Lake City, Utah 84112 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS<br>ARPA Order #2477 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Defense Advanced Research Projects Agency<br>1400 Wilson Blvd.<br>Arlington, Virginia 22209 | | 12. REPORT DATE<br>June 1976<br>13. NUMBER OF PAGES<br>130 |
| 14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)* | | 15. SECURITY CLASS. *(of this report)*<br>UNCLASSIFIED<br>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT *(of this Report)*

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

3-D, human color vision, neurophysiological, psychophysical, modelling neural, color, digital images, masking techniques,Shannon's theory, perceptual space, distortion, coding experiment.

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

A three-dimensional homomorphic model of human color vision based on neurophysiological and psychophysical evidence is presented. This model permits the quantitative definition of perceptually important parameters such as brightness, saturation, hue and strength. By modelling neural interaction in the human visual system as three linear filters operating on perceptual quantities, this model accounts for the automatic gain control properties of the eye and for the automatic gain control properties of the eye and for brightness and color (con't)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

20. ABSTRACT (con't)

contrast effects.

In relation to color contrast effects, a psychophysical experiment was performed. It utilized a high quality color television monitor driven by a general purpose digital computer. This experiment, based on the cancellation by human subjects of simultaneous color contrast illusions, allowed the measurement of the low spatial frequency part of the frequency responses of the filters operating on the two chromatic channels of the human visual system. The experiment is described and its results are discussed.

Next, the model is shown to provide a suitable framework in which to perform digital images processing tasks. First, applications to color image enhancement are presented and discussed in relation to photographic masking techniques and to the handling of digital color images. Second, application of the model to the definition of a distortion measure between color images (in the sense of Shannon's rate-distortion theory), meaningful in terms of human evaluation, is shown. Mathematical norms in the "perceptual" space defined by the model are used to evaluate quantitatively the amount of subjective distortion present in artificially distorted color images. Third, applications to image transmission and coding are presented. Results of a coding experiment yielding digital color images coded at an average bit rate of 1 bit/pixel are shown.

Finally conclusions are drawn about the implications of this research from the standpoint of psychophysics and of digital image processing.